AD 786667

AD-786 667

OBJECTIVE MEASUREMENT TECHNIQUES FOR EVALUATING VOICE COMMUNICATION CHANNELS

R. W. Hubbard, et al

Office of Telecommunications

Prepared for:

Federal Aviation Administration

July 1974

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

CREDIT STATEMENT

Some of the figures (identified) in this report are reproduced with permission from the American National Standard, S3.5-1969, copyright 1969, by the American National Standards Institute, copies of which may be purchased from the American National Standards Institute at 1400 Broadway, New York, NY 10018.

Technical Report Documentation Page

1. Report No. FAA-RD-74-77	2. Government Accession No.	3. Recipient's Catalog No. AD 786 667
4. Title and Subtitle OBJECTIVE MEASUREMENT TECHNIQUES FOR EVALUATING VOICE COMMUNICATION CHANNELS		5. Report Date July 1974
		6. Performing Organization Code
7. Author(s) R.W. Hubbard, W.J. Hartman		8. Performing Organization Report No.
9. Performing Organization Name and Address Institute for Telecommunication Sciences Office of Telecommunications U.S. Dept. of Commerce Boulder, Colorado 80302		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. DOT-FA72WAI-326
12. Sponsoring Agency Name and Address Department of Transportation Federal Aviation Administration Systems Research and Development Service Washington, DC 20590		13. Type of Report and Period Covered Final Report - Task A
		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This report presents methods for objectively evaluating the quality of voice communication circuits, based upon the Articulation Index (AI). The Speech Communication Index Meter (SCIM) and associated computer programs for calculating AI are discussed in some detail and procedures are presented for making system measurements.

17. Key Words Articulation Index; Performance Testing; Subjective & Objective Scoring; Voice Communication	18. Distribution Statement Document is avilable to the public through the National Technical Information Service, Springfield, VA 22151

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages [illegible]	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

OBJECTIVE MEASUREMENT TECHNIQUES FOR EVALUATING VOICE COMMUNICATION CHANNELS

R.W. Hubbard*
W.J. Hartman*

This report presents methods for objectively evaluating the quality of voice communication circuits, based upon the Articulation Index (AI). The Speech Communication Index Meter (SCIM) and associated computer programs for calculating AI are discussed in some detail and procedures are presented for making system measurements.

1. INTRODUCTION

There has long been a need for an inexpensive and reliable method to evaluate the quality of voice communication systems. The requirement is certainly evident when we consider the fact that very few (if any) voice communication systems are ever procured on the basis of performance delivered to the user; i.e., the quality and intelligibility of the speech at the ear of the listener. Most system specifications stop short of this requirement and rely on some engineering parameter such as signal-to-noise ratio (S/N) to serve as the final requirement.

*The authors are with the U.S. Department of Commerce, Office of Telecommunications, Institute for Telecommunication Sciences, Boulder, Colorado 80302.

In operating systems, measures of performance are also based upon engineering parameters in most cases. User performance tests are sometimes specified using subjective procedures that involve trained speakers and listening panels to directly score intelligibility. These tests are generally expensive and time consuming to perform, and the results are subject to statistical variation that must be understood. For example, it has been shown [Egan, 1948] in subjective scoring that performance evaluations differ with the following factors:

a) talkers used in the tests,

b) listeners, including their state of training for the tests,

c) sound levels used in scoring,

d) scoring techniques,

e) ambient background noise,

f) vocabulary used in the tests,

g) degree of familiarity with both the system and vocabulary on the part of talkers and listeners.

Many subjective evaluation procedures have been devised, however, and with consideration of the above factors can be quite effective [Stuckey, 1963] when properly administered. These procedures are based on a number of different vocabularies, including complete sentences, phonetically

balanced (PB) word lists, rhyme tests, and nonsense syllables.

As stated previously, subjective evaluations are time consuming and expensive to perform. Thus, they are not convenient to apply each time the performance of a new or existing system is to be measured. In addition, they are not always the best methods to apply when evaluating complete system performance, as found by Dabbs and Schmidt [1972]. An objective technique for evaluation would be preferable; particularly one that can be well correlated with subjective scoring, and that can be applied inexpensively in many laboratories with comparable results. The objective method should also be one that relates closely to other system parameters for design and analysis use.

It is the purpose of this report to discuss an objective performance measure known as the Articulation Index (AI), and to present the methodology of measuring and evaluating this parameter as an objective technique to score the quality of a voice system or channel.

The performance criteria for any voice communication system should ultimately be based upon intelligibility defined by the system users. Therefore, any useful objective performance measure should have a high degree of correlation with subjective tests. This report concentrates

on this aspect of using the AI objective method, and presents some specific limitations and some correction data that can be used in applying the method.

In addition to the AI method, the report includes a brief discussion of other methods based upon spectral and time-domain processing of the speech signal.

2. OBJECTIVE PERFORMANCE MEASURES

2.1 The Articulation Index (AI)

An objective scoring technique that has been developed and well documented for voice communication systems, is the computation of the Articulation Index (AI) [French and Steinberg, 1947; Kryter, 1962a]. The AI is a physical measure that has been shown to be highly correlated under certain conditions with the intelligibility of speech as evaluated by speech perception tests administered with groups of talkers and listeners. The calculation methods have been drafted in standard form by the American National Standards Institute, Inc. [ANSI, 1969], and the index itself is calculated from acoustical measurements (or estimates of the speech spectrum) and the effective masking spectrum of any noise which may be present along with the speech at the ear of a listener. The computation results in a weighted fraction representing, for a given speech channel and noise

condition, the effective proportion of the normal speech signal that is available to a listener.

The Articulation Index is considered to be one of the best objective techniques currently available for scoring the articulation of a voice system [Kryter, 1962b; Busch, 1969; Gierhart et al., 1970]. Comparison of the AI score with the average of several subjective listening tests for the male talker and test vocabularies is shown in figure 1. Comparisons of this type indicate the utility of the AI. If a performance index such as AI can be conveniently measured for a voice communication system, it may be related directly to an expected intelligibility for a given type of vocabulary. Success with such a technique would significantly reduce the requirement for direct subjective testing.

The work of French & Steinberg [1947] which led to the formulation of the AI method, is based upon the degrading effects of noise power over the voice spectrum range, including spread-of-masking. Spread-of-masking is a term used to note the degradation of speech intelligibility in a given spectral region due to noise or interference that exists in some other region of the speech spectrum. The intelligibility of a speech signal was found to be proportional to the signal-to-noise ratio averaged over 20

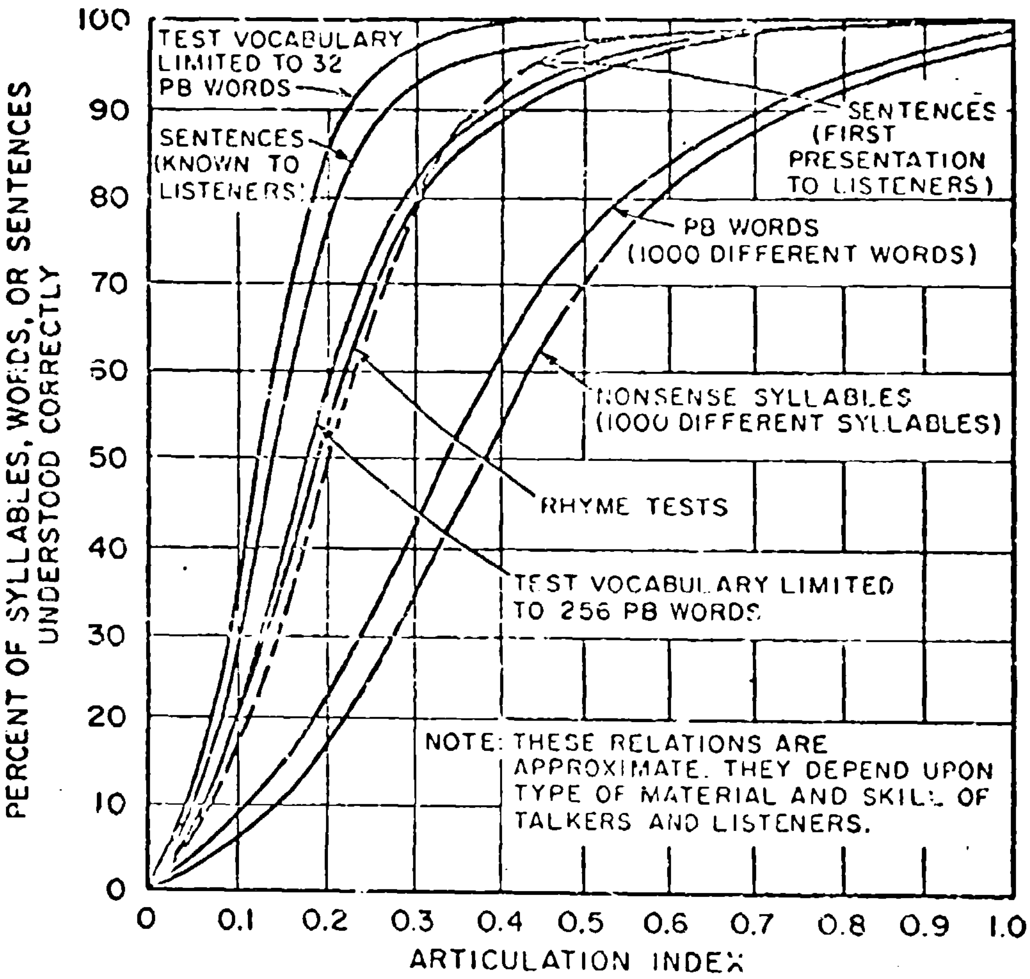


Figure 1. Relation between AI and various measures of speech intelligibility in the presence of white Gaussian noise (ANSI, 1969).

specified bands of frequency. These 20 bands were found to contribute equally to speech intelligibility. They are listed in table 1.

The computation of AI as specified by the ANSI [1969] Standard is a method that requires the manual plotting of spectral data, and a step-by-step procedure for evaluating

Table 1. Frequency bands of equal contribution to speech intelligibility, American National Standard (1969)

Band No.	Limits (Hz)	Mid-Frequency (Hz)
1	200-330	270
2	330-430	380
3	430-560	490
4	560-700	630
5	700-840	770
6	840-1000	920
7	1000-1150	1070
8	1150-1310	1230
9	1310-1480	1400
10	1480-1660	1570
11	1660-1830	1740
12	1830-2020	1920
13	2020-2240	2130
14	2240-2500	2370
15	2500-2820	2660
16	2820-3200	3000
17	3200-3650	3400
18	3650-4250	3950
19	4250-5050	4650
20	5050-6100	5600

the signal-to-noise power ratios in several spectral regions of the speech signal (20 bands, 1/3 octave bands, or full octave bands). These ratios are then summed over all bands and normalized by a constant multiplied by the number of bands used in the computation. This results in a fraction defined as the AI (≤1.0), and for the 20 band method is given by

$$AI = \frac{1}{600} \sum_{i=1}^{20} [10 \log_{10} s_i/n_i + 12 \text{ dB}], \qquad (1)$$

where s_i = speech signal power in band i,

n_i = noise power (including spread-of-masking) in band i,

a $S_i/N_i = 10 \log_{10} s_i/n_i$ dB,

$-12 \text{ dB} \leq S_i/N_i \leq 18 \text{ dB}$ by definition.

The limits placed upon the signal-to-noise ratios in (1) are based upon the research leading to the AI measurement. It was found that if a S_i/N_i for any given band (of equal contribution to intelligibility) exceeded approximately 18 dB, no further contribution from that band was discernable. Likewise, if the S_i/N_i degraded below a value of -12 dB no further loss in total intelligibility could be

detected. Thus the critical values based on intelligibility measures are considered to be bounded by these values.

The +12 dB term added to the S_i/N_i values in (1) is derived from considering the peak-to-average power ratio of a speech signal. It is shown, for example, by Gierhart et al. [1970], that the speech signal can be modelled with a modified Laplace distribution function. Theoretically, this function would have an infinite value of peak power. However, if one chooses a quasi-peak signal value such that actual peaks will exceed the chosen value with a probability of near 0.004, the peak-to-average power ratio is approximately 12 dB. The measured values of S_i/N_i in (1) are average values, and thus the +12 dB adequately accounts for the peak value of these ratios above the average. In addition, we note from (1) that this value also restricts the lower bound of AI to zero, if all the S_i/N_i values are -12 dB or less. This is a desirable condition because it directly relates an AI=0 to a complete lack of speech intelligibility. On the other end of the AI scale, we wish to have an AI=1 relate to near perfect intelligibility. Thus the normalizing factor in (1) is derived on the basis that if all 20 bands of analysis provide an $S_i/N_i \geq 18$ dB, the summation will result in a value of 20 x30 dB, or 600 dB.

The normalizing factor of 600 reduces this measure to unity, the condition for essentially perfect intelligibility.

2.2 Limitations & Problems Associated with AI

There are a number of limitations involved with the application of the AI technique to system measurements. Some of these are listed and discussed in the ANSI [1969] Standard, and must be understood before valid measurements can be obtained under the noted restrictions. It is the purpose of this section to discuss these known limitations and to outline the necessary corrections when the AI is measured under the noted conditions.

2.2.1 Spectrum considerations in the AI method

The AI measurement is based upon spectral considerations of the adult male voice. It has not been evaluated for application directly to the female voice or that of a child, which would generally contain higher frequencies in the spectral range than those of the adult male. A careful study of this aspect may require that an adjustment be made to the equal-bands-of-contribution to intelligibility from those shown in table 1. Such a study has not been made to our knowledge. However, the results of such a study would perhaps be insignificant in respect to results such as those of figure 1 due to general variability of subjective data. For example, subjective performance data usually involve

significant standard deviations from the mean due to the factors noted in the Introduction. Thus, the curves of figure 1 denote only the mean values of measurements made by a number of investigators, and represent the norm of the best subjective data available. It would seem appropriate to develop similar subjective data based on the female and/or childs voice, rather than to repeat the AI spectral-band analysis for these cases. In other words, it is felt that there is probably more variation in the subjective data scores than would be found in the spectral changes for the AI analysis bands. In any event, comparative data such as those in figure 1 for other than the male voice would be the criteria required.

Until calibration data are available for other than the adult male voice, the data of figure 1 can be used as a general guide in other situations. It must be kept in mind, however, that the relationship is not precisely known.

A second spectral consideration in AI measurements is that of sample time or the length of the speech signal from which a spectral estimate is made. The AI technique is based upon the long-term average spectrum of the male voice. This characteristic is noted in section 4. In order to perform a valid measurement, the spectrum of the test speech signal must be obtained over a sufficiently long sample to

approach the long-term spectral reference. For speech signals that are continuous in nature, the sample or averaging period should be on the order of 1 minute or longer. For a synthetic test signal, such as that used in instrumented AI systems (see section 3), the averaging time can be less as explained in a later discussion of these instruments.

Frequency distortions of the speech signal are also important in application of the AI performance measure. The AI Standard [ANSI, 1969] indicates that the method accounts for certain distortions due to unequal transmission gains, provided that the unequal emphasis is applicable to only one of the following at a time:

1. the high frequency components of the speech signal
2. the lower frequency components
3. the mid-frequencies

If distortions occur in combinations of the above, little is known about the validity of the AI. It is known, however, that the AI will not be valid in estimating intelligibility in cases where the speech spectrum is very irregular; i.e., when it is composed of peaks and valleys with slopes that exceed approximately 18 dB per octave. For these reasons, it will be important to monitor (in some manner) the spectrum of a received test signal when the AI method is

applied to communication system measurements. This requirement is discussed in a later section of this report.

2.2.2 Amplitude considerations in the AI method

Peak-clipping is frequently used in voice communication systems, in order to limit the peak-power required in transmitters and receivers where voice is the modulating signal. Licklider [1946] and others have shown that the intelligibility of speech is not seriously effected when the signal is subjected to rather severe degrees of peak-clipping. However, an amplitude distortion of this type changes both the time and frequency characteristics of the basic speech signal. Thus, if it is used in a system under test, it must be accounted for in the AI measurement. The ANSI Standard presents a correction method to be used, based upon prior knowledge of the level of clipping and the post-clipping amplification used in the system.

The usual peak-clipping process involves both of the above elements. For example, clipping the extreme peaks of a speech signal changes the amplitude probability distribution of the signal, and consequently, the rms level. An AI measurement made when both of these signal processes are used (clipping and amplification) will generally result in a slightly higher value than under normal conditions. The reason for this is that the S_i/N_i is slightly higher in

each band of analysis because of the higher rms signal level. This assumes that only the speech signal is effected by the clipping process, and not the additive noise or interference.

When measurements of AI are made in a speech-clipped channel, the results are generally reliable. If, however, it is desired to relate the measured value to that which would be found for an unclipped channel, a correction can be made. The details of the correction method are given in section 4.

2.2.3 Interference considerations in the AI method

The fundamental factor which causes degradation to the intelligibility of speech in a communication system is the additive type noise or interference. The AI method adequately accounts for this factor in the most common situations. However, there are certain interference conditions that must be corrected for in the AI measurement, and others where the effect is not precisly known. This section discusses these factors briefly, and section 4 presents some correction data that can be applied.

The AI method adequately treats the masking of steady-state wide-band Gaussian noise, and also those cases where the noise interference is at least 200 Hz in bandwidth anywhere in the range 200 to 6100 Hz [ANSI, 1969]. In cases

where the interference bandwidth is less than 200 Hz, which includes the problem of tone-type interference signals, the AI is known to be inaccurate. These latter cases have been studied, and the results as defined by the use of the SCIM (see section 3) system are given in section 4.

When the interfering noise is not steady-state, but is intermittent, the measured AI value is not accurate. However, it can be corrected if the time characteristics of the intermittent noise are known. An initial correction is made based on the duty cycle of the noise; i.e., the fraction of the time the noise is "on". The second correction is based upon the rate of the intermittent noise. The curves relevant to both of these conditions are given in section 4.

2.3 Other Developing Techniques

The AI measurement is the best known objective technique available at this time, and has been documented and standardized. For these reasons, it is considered to be the best method readily available to system users for performance testing. The instrumented techniques for performing the measurement in the form of SCIM have been applied for some time, with documented results. However, there are other methods in various stages of development that merit mention. They are not treated in any detail in

this report, but are noted for completeness and to alert the reader of possible new techniques.

2.3.1 Evaluation based on phonemes and spectra

Some work has been devoted to measuring intelligibility based on the probability of occurance and the recognition of phonemes in speech. An example is the development of a system called Correlation of Recognition of Degradation with Intelligibility Measurements (CORODIM) by the Philco-Ford Corporation[1]. It was based on a 300 word list of the consonant-vowel-consonant type. Comparisons are made in a computer scoring system that correlates stored phoneme information with similar data that are degraded in some manner during speech transmission. There is no known direct relationship between phoneme intelligibility and word intelligibility [private communication from Mr. Allen Busch, FAA, NAFEC]. However, the early reported[2] work on CORODIM showed good correlation between these factors. A recent inquiry [private communication from Mr. Charles Teacher, Philco-Ford Corp.] indicates that there has been little development on this system in the last few years.

---

[1] Willow Grove, Pa.
[2] Minutes of the Meeting of the Aeronautical Satellite Voice Processing and Subjective Testing Committee, April 17, 1970, COMSAT Laboratories, Clarksburg, Md.

Another new technique [reported by Teacher] based on the correlation properties of spectral data has been developed, and is currently under evaluation by the U.S. Air Force, Rome Air Development Center (RADC). It is called Automated Intelligibility Measurement (AIM). Evaluation data may be available from RADC in the near future, but no reports have been published to date.

2.3.2 Evaluation based on time domain processing

The Institute for Telecommunication Sciences (ITS) has been considering a technique involving speech processing in the time domain as a new objective scoring method. Current work is under the sponsorship of the Federal Aviation Administration (FAA). The method is based on the emerging technology known as Linear Predictive Coding (LPC) [Atal & Hanauer, 1971], from which methods of producing high-quality synthetic speech have been developed. The work on the scoring method has not progressed to the point that the technique can be described distinctly, but a technical report should be available from the ITS within a few months.

Processing in the time-domain appears to have several advantages over spectral processing. The time signature of a speech signal is more unique than spectral signatures, and the method should alleviate the problem of measuring and defining the noise or interference signals.

3. INSTRUMENTATION FOR AI MEASUREMENT

The measurement of the AI directly in a test instrument has been approximated in two independent developments. The first of these systems was developed in about 1960, and is used by the U.S. Army Electronic Proving Ground at Ft. Huachuca, Arizona. It is entitled "The Voice Interference Analysis System" (VIAS), and is described in a document prepared by Lockheed Electronics Co. [1972]. It is based on the AI process, but has a number of disadvantages when compared with a later development. For example, the VIAS test signal is composed of a triangularly modulated tone at 950 Hz. This signal is not representative of a speech spectrum and thus the weighting for the speech signal analysis is provided only in the analyzer. This technique does not directly allow for spectral distortions imposed by the system under test, or for time-variations in the transmission medium. In addition, the analyzer measures only the noise content over the analysis bands, and assumes a signal content based on the level of the 950 Hz tone. For these reasons, the VIAS system is not considered further in this report.

The second system developed for automatically measuring the AI is the Speech Communication Index Meter (SCIM). This development was sponsored by the U.S. Air Force, Electronic

Systems Command, under contract with Bolt, Beranek and Newman, Inc., of Cambridge, Mass. [Kryter & Ball, 1964]. The SCI measurement is patterned after the AI using a 9-band S/N measurement over the speech spectrum. The data are analyzed in an analog mode with an instrument that computes the S/N in each band, sums and normalizes the results, and provides a digital readout display of the decimal value of SCI. The SCI measure is similar to that given by (1), and can be written for the 9-band method as

$$SCI = \frac{1}{270} \sum_{i=1}^{9} [10 \log_{10} s_i/n_i + 12 \text{ dB}]. \qquad (2)$$

The 9 bands used in the SCI measurement technique are given in table 2.

The SCI measurement is obviously a compromise to the ANSI Standard for AI, but it has been demonstrated [Kryter & Ball, 1964] to be highly correlated with AI values calculated using (1).

The SCIM system includes a signal generator that develops a speech-shaped noise spectrum to simulate the voice signal, as well as several tones for calibrating the test system, and timing signals used to control the operation of the SCI analyzer. Instrumentation of this type

Table 2. SCIM analysis bands

Band No.	Limits (Hz)	Bandwidth (Hz)	AI Bands (Approx.)
1	250-500	250	1,2,3
2	500-750	250	4
3	750-1000	250	5,6
4	1000-1250	250	7
5	1250-1650	400	8,9,10
6	1650-2050	400	11,12
7	2050-2450	400	13,14
8	2450-3200	750	15,16
9	3200-4200	1000	17,18

provides a capability of performing on-line objective measurements of system performance.

Figure 2 illustrates the composition of the SCIM test signal. A 1-kHz calibration tone is provided to adjust proper signal levels for system testing. The first segment of the test signal is a speech-shaped noise signal having the same spectrum as the long-term average for a male talker. A 5-Hz square wave modulation is applied to the noise signal to simulate the natural pauses in human speech.

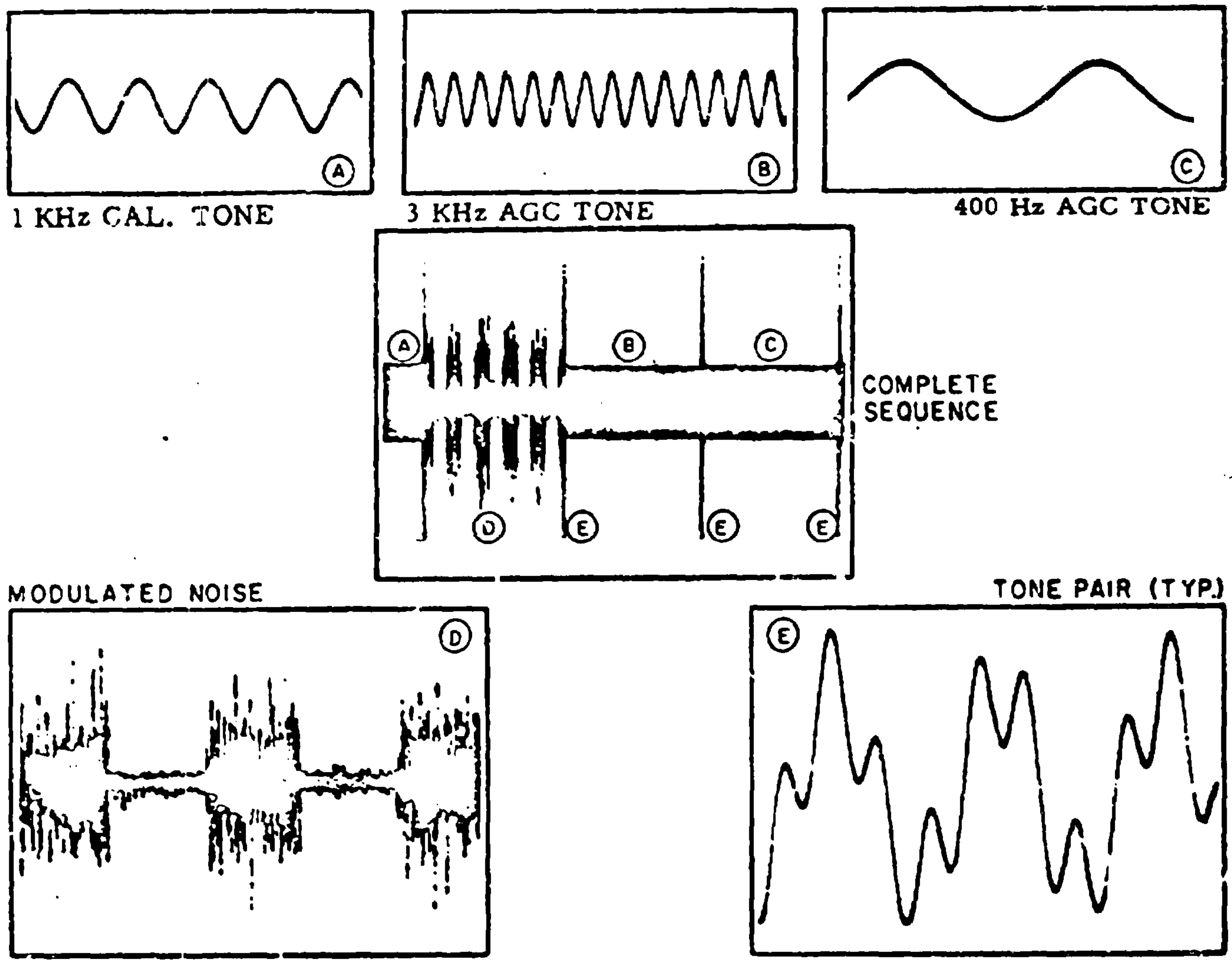


Figure 2. The SCIM test signal.

Following the speech-shaped noise, two sinuscidal tones are generated in sequence at 3 kHz and 400 Hz. These signals are generated at the same level as the calibration tone, and are used to maintain the signal level and thus stabilize any agc circuitry in the test system during measurement. The reason for using two tones for this purpose is explained below in the description of the SCIM analyzer.

Between each segment of the test signal a short synchronizing signal is generated, composed of a tone-pair burst of 2 kHz and 600 Hz. The sync signal is used to control the operation of the analyzer. The analyzer is an analog system designed to compute the S_i/N_i for each of the 9 frequency bands, limit the values as noted with (1), and compute the SCI score given by (2). The analyzer sequence is illustrated in figure 3. The signal power (s_i) in each band is computed during the "speech" portion of the test signal and stored in integrators. The noise power (n_i) in the first five spectral bands is computed during the 3 kHz tone portion. Note from table 2 that these bands are well below the 3-kHz tone. Similarly, the noise power in bands 6 through 9 is computed during the 400 Hz tone portion of the signal. The analyzer then combines these integrated values of signal and noise power, computes the ratios and sums them in accord with (2).

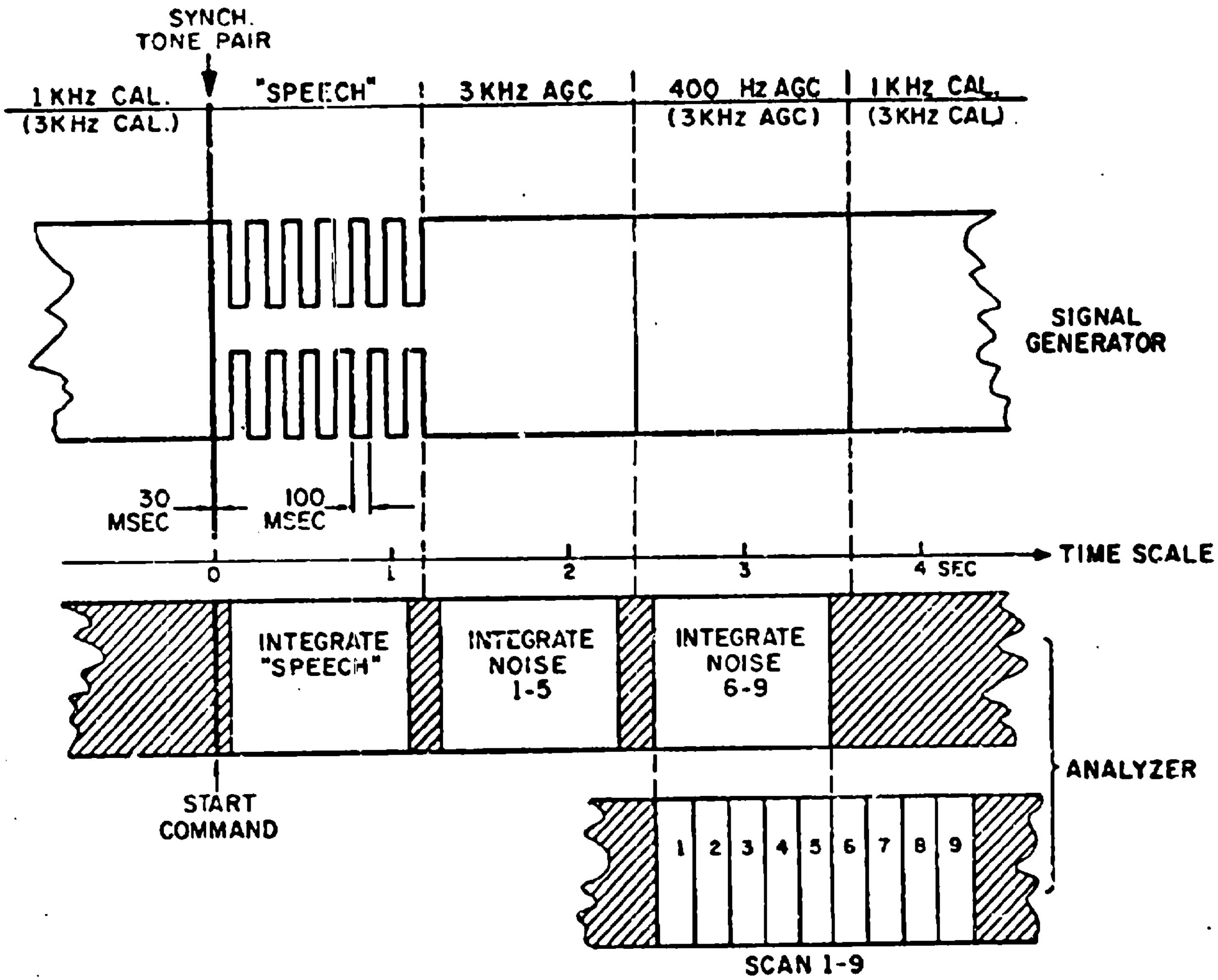


Figure 3. Timing sequence of SCIM signal generator and analyzer.

A simplified block diagram of the SCIM signal generator is shown in figure 4. A random noise source that generates a broad-band uniform spectral density is used as the basic signal generator. A spectrum shaping filter is used to weight the noise spectrum to that of the long-term average value for a male voice. The 5 Hz square wave modulation is applied by the modulator at a level such that the peak-to-null ratio is on the order of 16.5 dB. The commutating switch is driven by a timing circuit such that it dwells on the speech-shaped noise signal, the 3 kHz tone and the 400 Hz for approximately 1 sec during each signal sequence. The switch passes through the sync generator positions in a few msec, inserting the sync burst between the other signal elements. The signal sequence must be initiated by the operator from a front-panel start switch. The commutator switch begins and returns to the contact for the 1 kHz oscillator, providing this signal at the output for leveling and calibration procedures.

A functional block diagram of the SCIM analyzer is shown in figure 5. The test signal is divided in spectrum into the 9 analysis bands by the bank of bandpass filters and detectors. The signal and noise power levels are detected by the integrator banks during the respective portions of the test signal. Switching of the integrators between

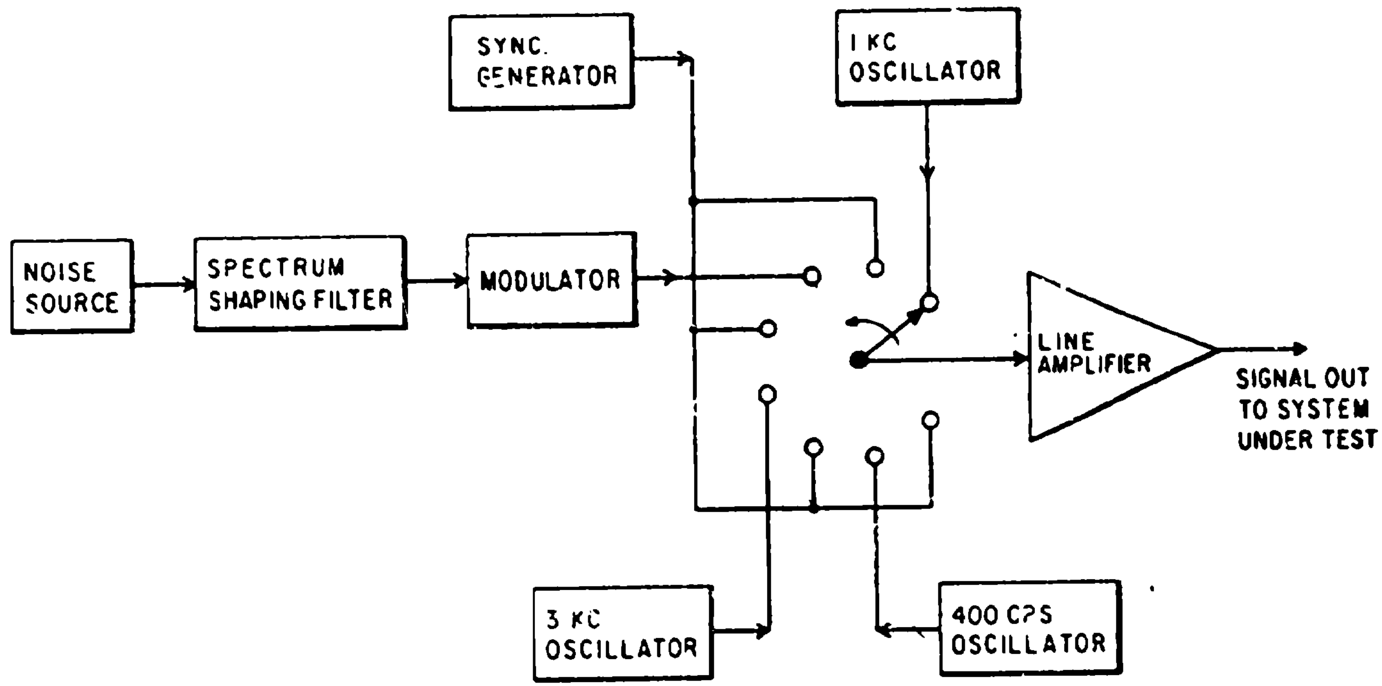


Figure 4. Simplified block diagram of the SCIM signal generator.

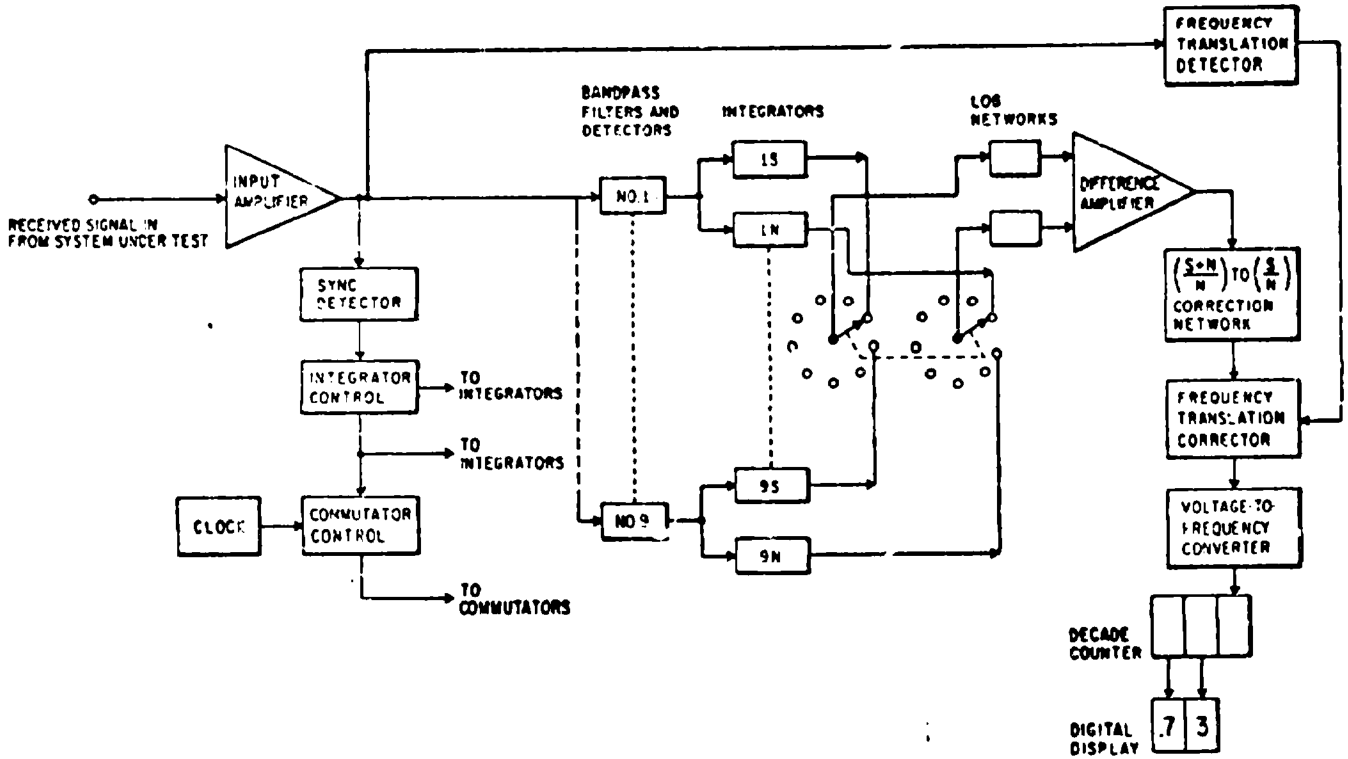


Figure 5   Simplified block diagram of the SCIM signal analyzer.

portions is accomplished through the sync detector circuitry. After integration of both the signal (plus noise) and noise power, the commutator control circuitry starts a scanning process of each of the 9 signal and noise integrators. Each of these values is fed in turn into the difference amplifier after conversion to logarithmic units. The difference amplifier is used to derive the S+N/N from the relation

$$\log \frac{(s+n)}{n} = \log(s+n) - \log n, \qquad (3)$$

expressed in dB. A correction network is used to convert (s+n)/n to s/n, based on the assumption that the interference is independent or uncorrelated with the signal. In this case, the s/n can be determined from

$$\frac{s}{n} = \frac{s+n}{n} - 1 \qquad (4)$$

The limits required in (1) are set by the difference amplifier and the correction network. The output signal of these stages then represents the analog solution of (2), and the voltage-to-frequency converter is used to provide a digital display of the computed SCI value.

Figure 5 also shows a frequency translation detector and correction circuit. The purpose of this part of the analyzer is to correct for any significant spectral shifts in the transmission system under test. The translation detector monitors the 1 kHz calibrate signal, and corrects the AI reading automatically if this signal has shifted in spectrum by more than about 50 Hz.

Spread-of-masking is an important factor in AI scoring. This term is used to describe the phenomenon that an interfering signal or noise at a given frequency can cause degradation to a speech signal at higher frequencies. The method used in the SCIM process to account for this factor is discussed further in the Appendix.

The SCIM instrumentation is relatively expensive, and only a few systems are currently known to be in operation. Those that do exist are rarely used in on-line measurements, and data are usually tape recorded in field tests and analyzed with the use of the SCIM analyzer in the laboratory.

In order to provide an analysis capability to organizations and laboratories that do not have a SCIM facility, off-line procedures have been developed by ITS. The procedures use computer programs to analyze a SCIM signal [Hubbard & Payne, 1974]. Two such methods have been

developed under the sponsorship of the FAA and the National Aeronautics and Space Administration (NASA), Goddard Space Flight Center (GSFC). Each of the computer methods is discussed in the Appendix.

The original design of the SCIM Signal Generator and Analyzer are shown in figures 6 and 7 respectively. A new integrated circuit version of this equipment has been considered by the manufacturer[3], but is not commercially available at this time. A signal generator patterned after the SCIM generator has been developed in the ITS Laboratories, using integrated circuitry and solid-state switching [Payne & McManamon, 1973]. This generator was designed primarily for applications involving acoustic coupling to telephone networks, and off-line data processing using the computer methods noted above. In this application, the sync signal is not necessary and thus is not provided in the ITS generator.

4. SCIM MEASUREMENT PROCEDURES

Section 2 of this report has noted some limiting considerations in obtaining valid AI measurements. Section 3 has discussed the instrumentation for the SCIM method of measuring AI. In this section, we present the data necessary for making corrections to measured values (where

[3] Bolt, Beranek & Newman, Inc., Cambridge, Mass.

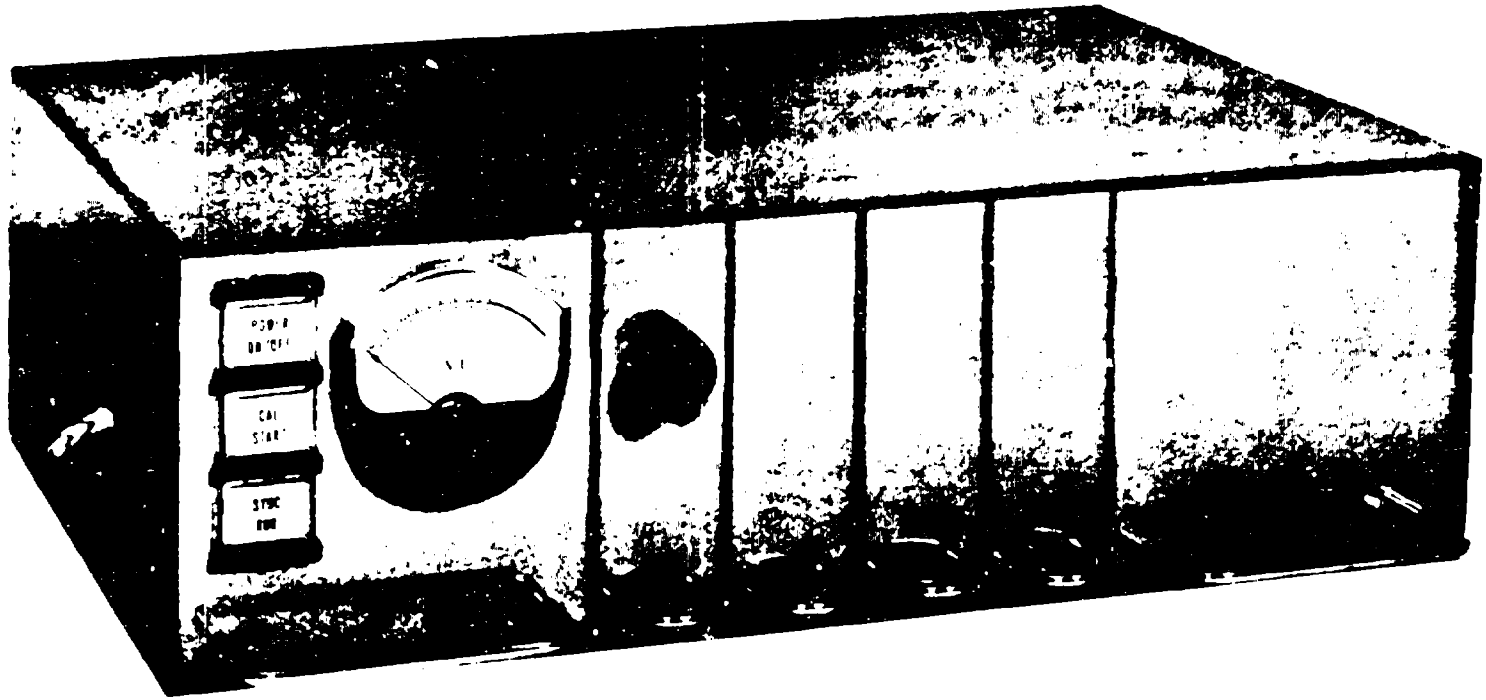


Figure 6. SCIM signal generator.

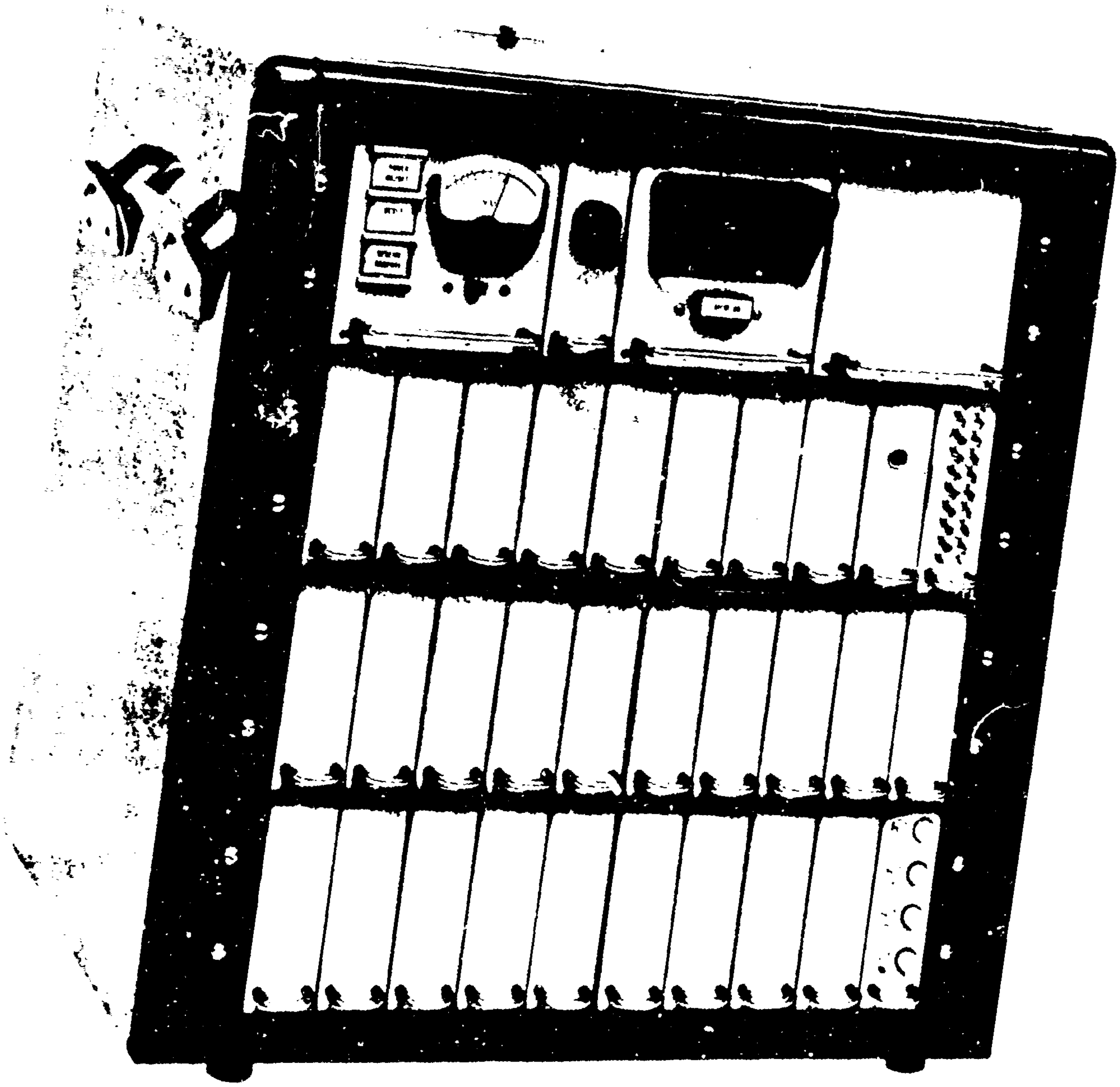


Figure 7. SCIM analyzer.

required) and the results of other effects on the SCIM measurement determined from laboratory studies.

4.1 Corrections to SCI Values

The three considerations of spectrum, amplitude and interference effects on AI noted in section 2 are discussed in this section. The specific correction data involved are presented, as well as other information developed for improving the reliability of scores measured using SCIM techniques. The following discussions will be relevant in most cases only to the analog voice channel. Specific problems encountered in the digital voice processes are noted in section 4.2.

4.1.1 Spectral monitoring & correction

The SCIM analysis is based upon the long-term average spectrum of the male voice which is shown in figure 8. For the actual voice signal the spectrum must be measured over a minute or longer of continuous speech. However, for the SCIM system where a synthetic speech spectrum is derived from a random noise signal, good spectral estimates can be made during samples as short as 1 sec. Both the SCIM analyzer and the computer scoring methods noted in the Appendix yield valid results from the standard SCIM signal, and no correction is required for the sample length.

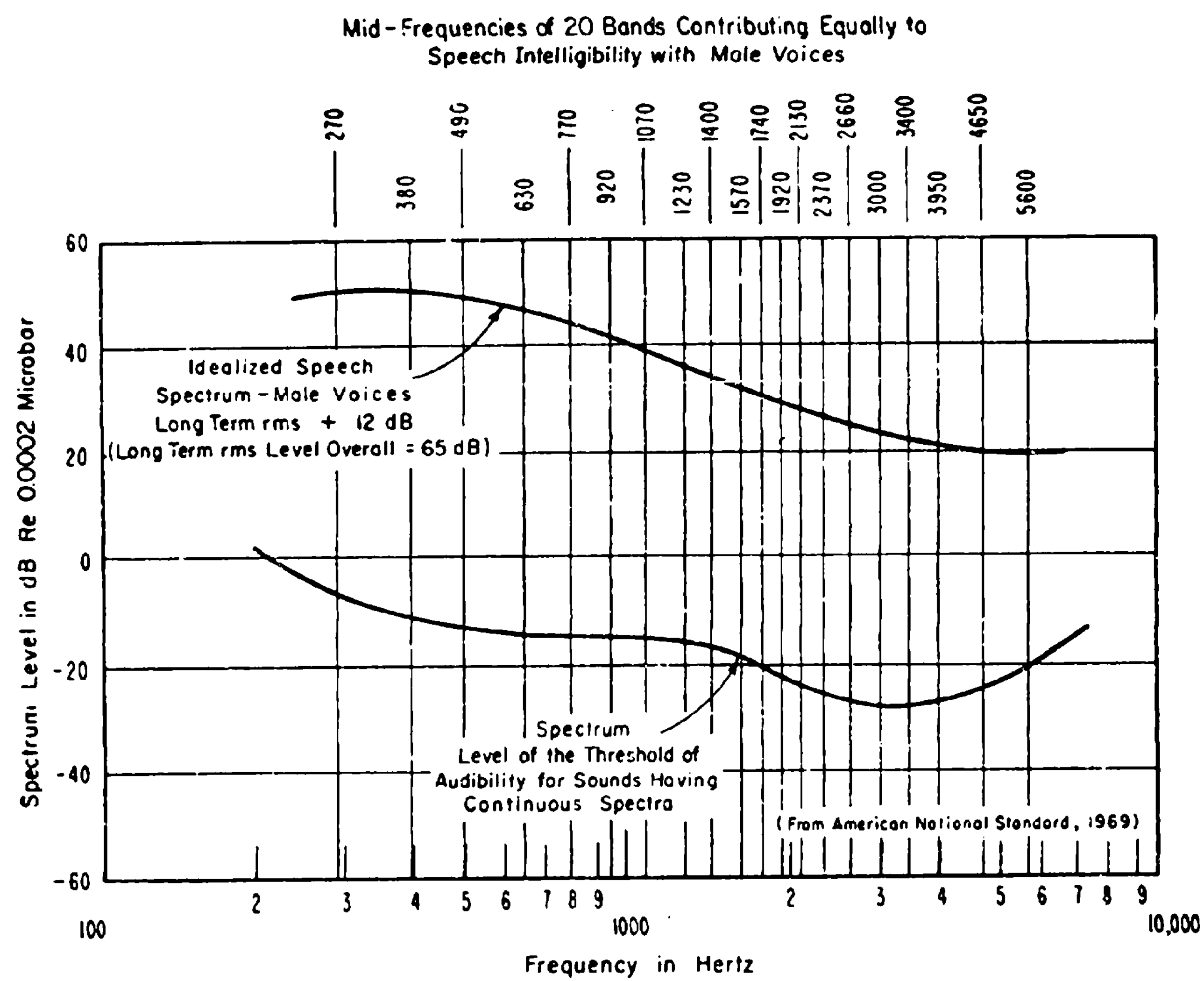


Figure 8. Long-term average spectrum of male speech.

Frequency distortions, however, that may be present in the analyzed signal, must be accounted for. Using the SCIM analyzer directly in on-line tests poses the most difficult problem. For example, an on-line spectrum analyzer should be used to measure the received test signal to directly determine the presence of spectral distortion. In cases where frequency selective distortions are present, the character of the spectrum can be noted from the analyzer and compared with the notations in section 2.2.1 to determine the validity of the SCI measure. Where a spectrum analyzer is not readily available, the SCIM analyzer (when properly aligned) can be used to derive a gross spectrum signature of the test signal. This is accomplished by observing on an oscilloscope the dc voltage input to the voltage-to-frequency converter of the analyzer. This signal will provide a bar-graph display of the 9-band integrator signal outputs as the commutater switch scans these circuits. When the analyzer is aligned properly, and no appreciable spectral distortion is present, the display will be 9-segments at essentially equal voltage levels. Distortion present in any of the 9-bands will be evident from a dc offset of the bar for the band(s) in which the distortion occurs. There is no problem in determining the spectral

distortion from either of the computer scoring methods, as both start with spectral analysis results.

In order to assess the relative spectral importance of the 9-bands in the SCIM system measurement, tests were made in which each of the 9-bands were filtered out of the test signal before analysis. The results are shown in figure 9. Based on the theory of equal intelligibility contribution for each band, the expected result is indicated by the dashed line in the figure. It can be seen that the result does have a median corresponding roughly to the expected value. It will be noted, however, that eliminating the spectral energy in bands 3 and 4 has the most significant effect on the measured value. These data were developed using digital filtering techniques where extremely steep filter skirts could be achieved.

Figure 10 shows the integrated results of continuing the band elimination tests as that for figure 9, but extending the single-band process to combinations of two and three contiguous bands for each test. The curve of figure 10 is the mean of these tests and is the best estimate possible of the static spectral distortion effects on SCIM. Note that the curve indicates the largest loss in score is caused by distortions at the low frequency end of the spectrum. It is suggested that the data of figure 9 be used to correct SCIM

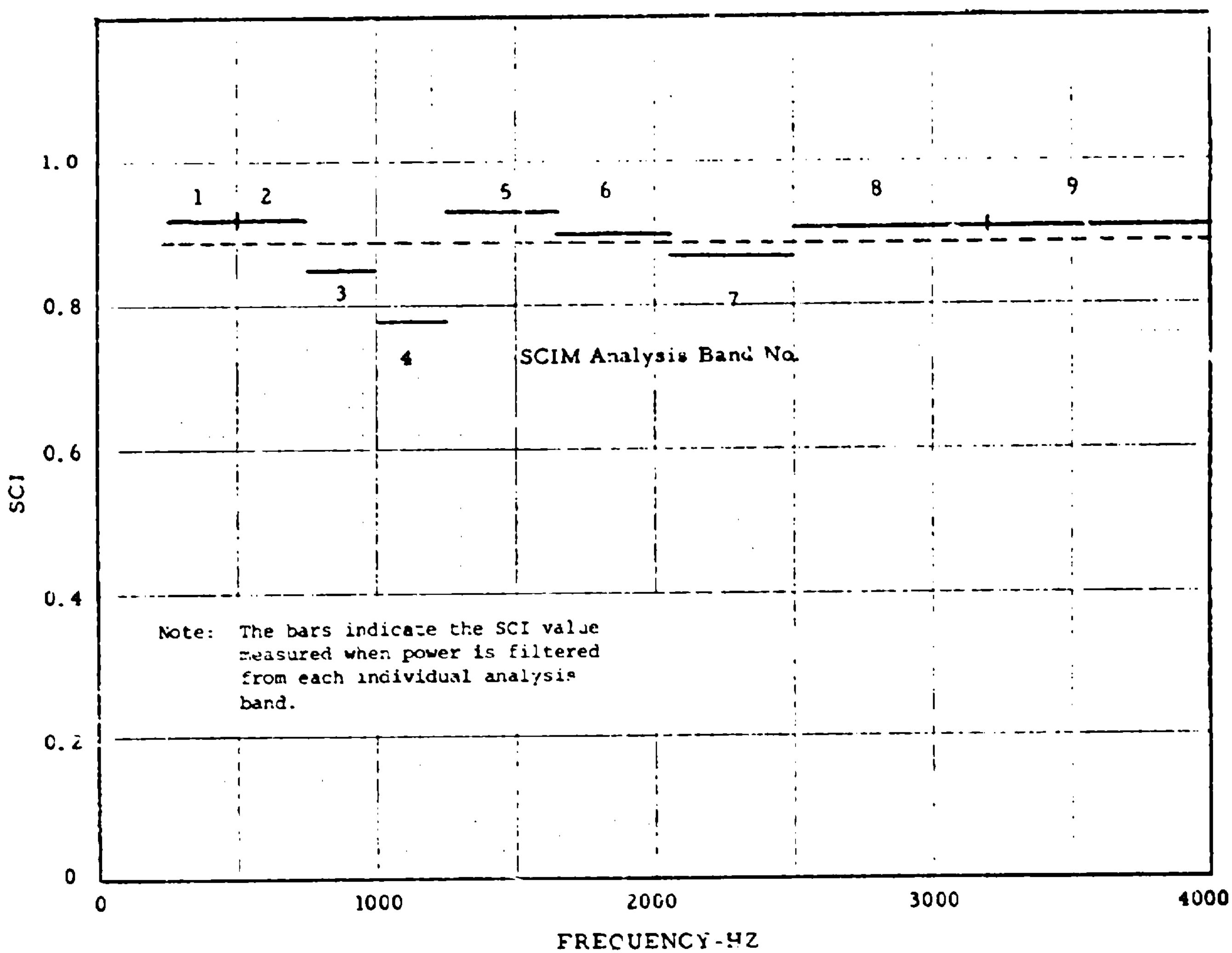


Figure 9. Band elimination test on SCIM system.

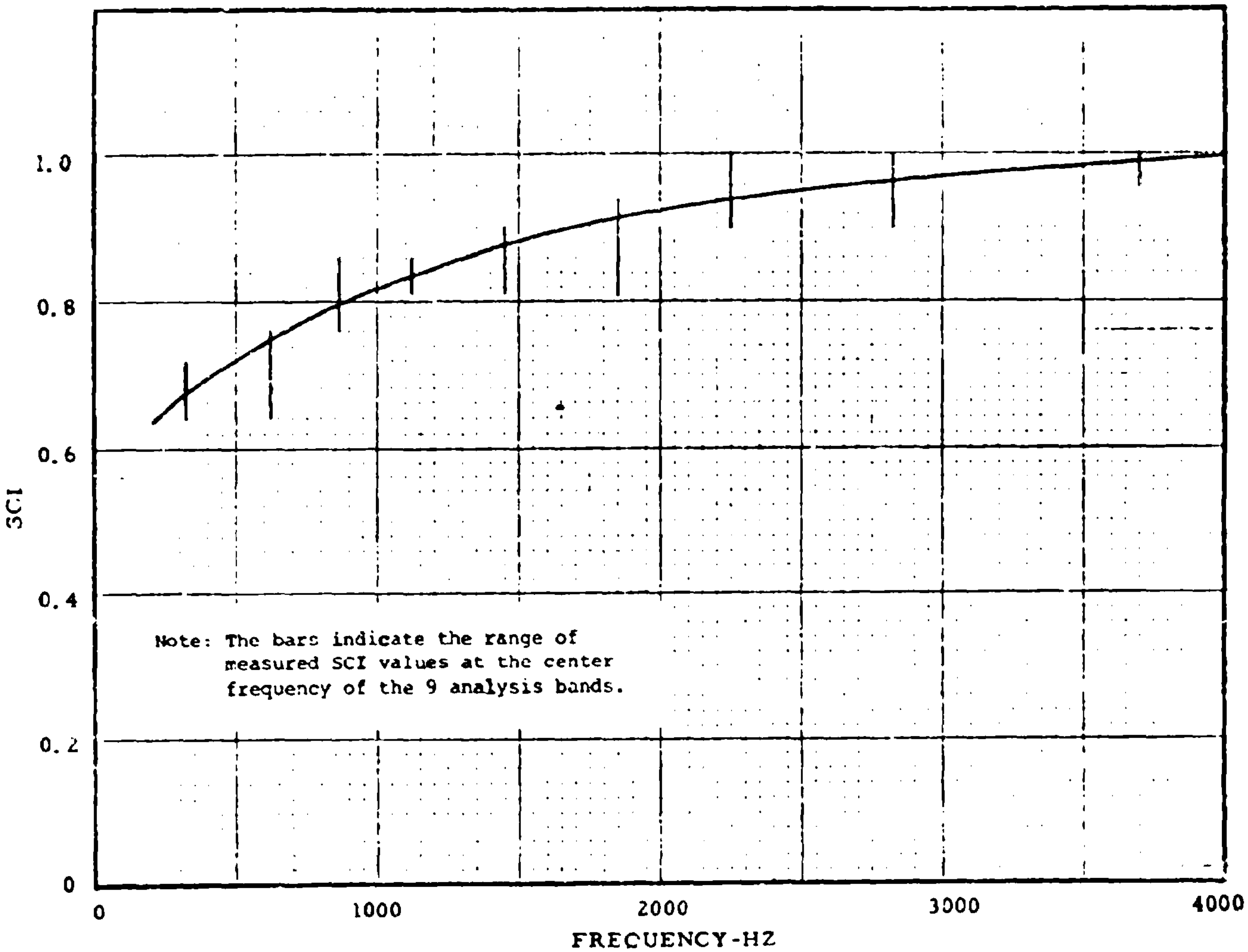


Figure 10. Multiple contiguous band elimination tests on SCIM.

readings when the distortions in the signal spectrum are on the order of 250 Hz wide. For distortions between 250 Hz and approximately 700 Hz, figure 10 should provide a good estimate for correction. These results are based, of course, on complete spectral elimination. For less severe spectral losses, the correction should be scaled to the degree of estimated loss based on a linear power scale. In any case, the low frequency correction is the most significant to be made.

Using figures 9 & 10 in conjunction with the spectral estimates noted above, adjustments can be made in the observed SCIM score. Associated subjective data are needed to completely calibrate these effects; these data have not yet been evaluated. A cursory subjective evaluation has been made, however, to indicate that severe spectral distortions can be tolerated in a static sense without great loss in intelligibility. Voice quality and recognition are lost in severe cases, but intelligibility tends to remain high.

4.1.2 Corrections for speech clipping

SCIM measurements made in a speech-clipped channel will generally be higher than those measured in non-clipped systems. The data presented here permits a correction

factor to be determined from a speech-clipped measurement that will score the standard voice channel.

Figure 11 illustrates the increase in the long-term spectral level of peak-clipped speech as a function of the sum of the peak-clipping level and post-clipping amplification. If these values are known for the system under test, the appropriate ordinate value (in dB) can be used to compute a corrected SCI score for an unclipped measurement. For example, the characteristic in figure 11 indicates that for a clipping level plus post-clipping amplification of 6 dB (3 dB each), the spectral power in the test signal is increased 5 dB in each of the spectral analysis bands. Equation (2) for this difference in signal power would yield an SCI value 45/270=0.17 units higher than would be measured in the unclipped speech channel. If we let P_i be the power level (dB) increase in the clipped-speech channel, then the correction for an unclipped channel can be computed from

$$\Delta SCI = - 1/270 \, [9 \, P_i]. \qquad (5)$$

In cases where the level of clipping and the post-clipping amplification are not equal, the difference between these factors must be subtracted from the value of P_i used in (5).

NOTE: This figure shows the increase in rms speech level as a function of clipping when clipped level is raised to clipping reference level.

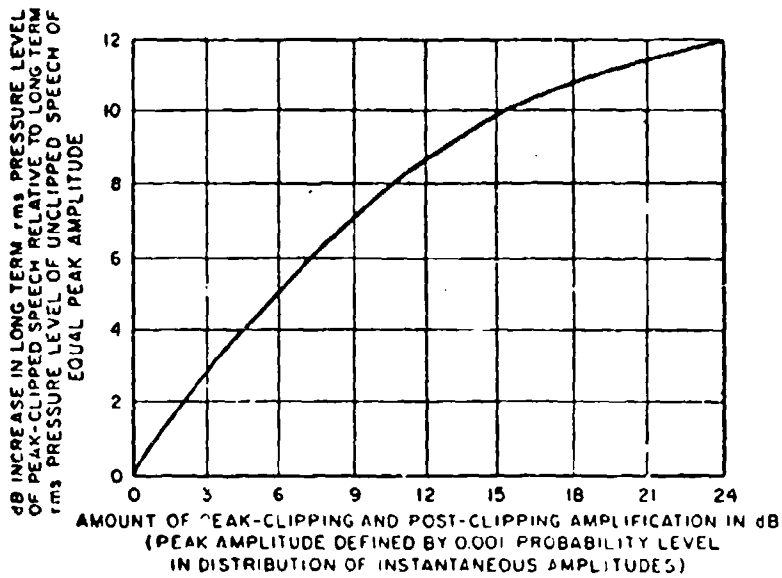


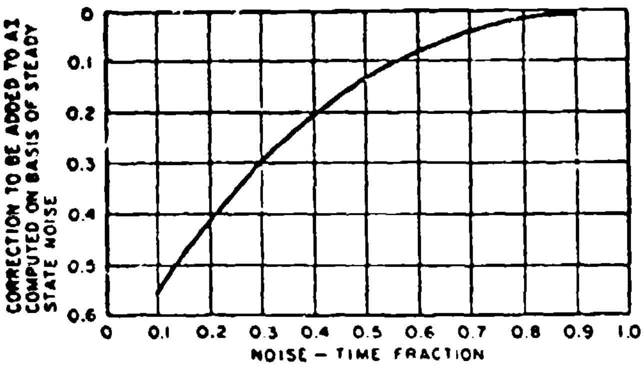
Figure 11. Speech level as a function of peak clipping (ANSI, 1969).

4.1.3 Corrections for non-steady-state noise

As noted in section 2, the measurement of an SCI value will be accurate only if the masking noise is steady-state in character. When the masking noise has a definite duty cycle and/or is interrupted in time, a correction to the observed SCI value is required. Figure 12 shows the correction to be made for the duty cycle (noise-time fraction). For example, if the masking noise can be observed to have a duty cycle of 0.3 (noise "on" for 30% of the base period and "off" 70%), figure 12 indicates that a

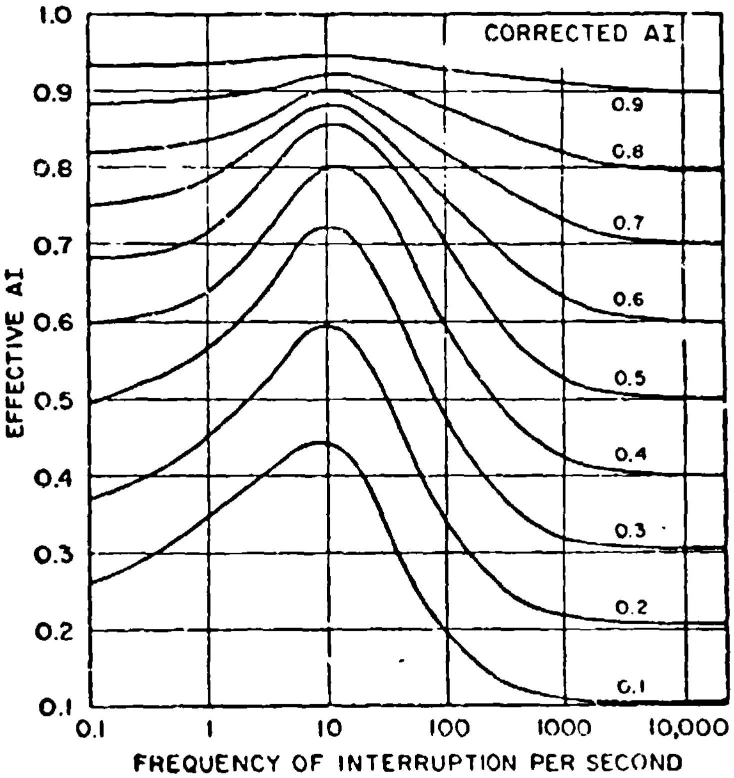
+0.3 correction be made to the measured value. In addition to this correction, another may be required if the repetition rate of the masking noise is less than approximately 1000 Hz. The second correction is made with use of the curves in figure 13. Assume that an SCI=0.3 value has been measured in the above example before correction. The corrected value from figure 12 is (0.3+0.3)=0.6. Entering the curve for this value in figure 13, and assuming an interruption rate of 100 Hz, we find the effective SCI value to be 0.76. These data are from the ANSI [1969] Standard, and are after the work of Miller and Licklider [1950].

In order to make these corrections to system measurements, some means of monitoring the actual time signatures of the masking noise must be provided. For on-line measurements using SCIM, this means that an oscilloscope analysis of the system noise should be made. In the case of off-line measurements, the character of the interruptive or masking noise can be made from the calibration tone portion of the SCIM test signal. This scheme is readily implemented using the TD/100 system discussed in the computer routine of the Appendix.



NOTE: The ordinate shows a correction to be applied to the articulation index computed on the assumption that a masking noise is steady-state for various noise-time fractions. The corrected AI cannot exceed 1.0.

Figure 12. Correction for various noise-time factors (ANSI, 1969).



NOTE: This figure shows the effective AI as a function of the frequency with which a masking noise is interrupted. The parameter of the curves is the corrected AI calculated on the assumption that the masking noise is steady-state and then adjusted according to Fig. 9 for the fraction of the time the noise is on.

Figure 13. Effective AI as a function of interruption rate (ANSI, 1969).

4.1.4 Effects of noise bandwidth and tone interference

The SCIM system was tested for its response to both random noise interference of varying bandwidth, and to single frequency tones within the speech bandwidth. The results of these tests are summarized in figures 14 and 15, respectively. It is seen from figure 14, that noise bandwidths up to the order of 100 Hz have no significant effect on the measured SCI value for total S/N values greater than 3 dB. Note also that the 20 dB curve in this figure at a noise bandwidth of 4 kHz corresponds to the theoretical situation for an SCI value of 0.5. The measurements indicate close agreement to the theoretical value. The effects of noise bandwidths between 100 Hz and 4 kHz on subjective scores are not yet known. The subjective data associated with these tests have not been scored.

A similar set of measurements using tone interference were also made. The results are shown in figure 15 for a number of sinusoidal signal power levels relative to the power in the "speech" portion of the SCIM test signal. Note that the effects of these interference situations are quite similar as far as the SCIM system is concerned. However, it is known that tone interference can be much more serious in subjective tests than broad-band noise. Here, again, the corresponding subjective material has not yet been scored to determine the tone interference effect on actual voice

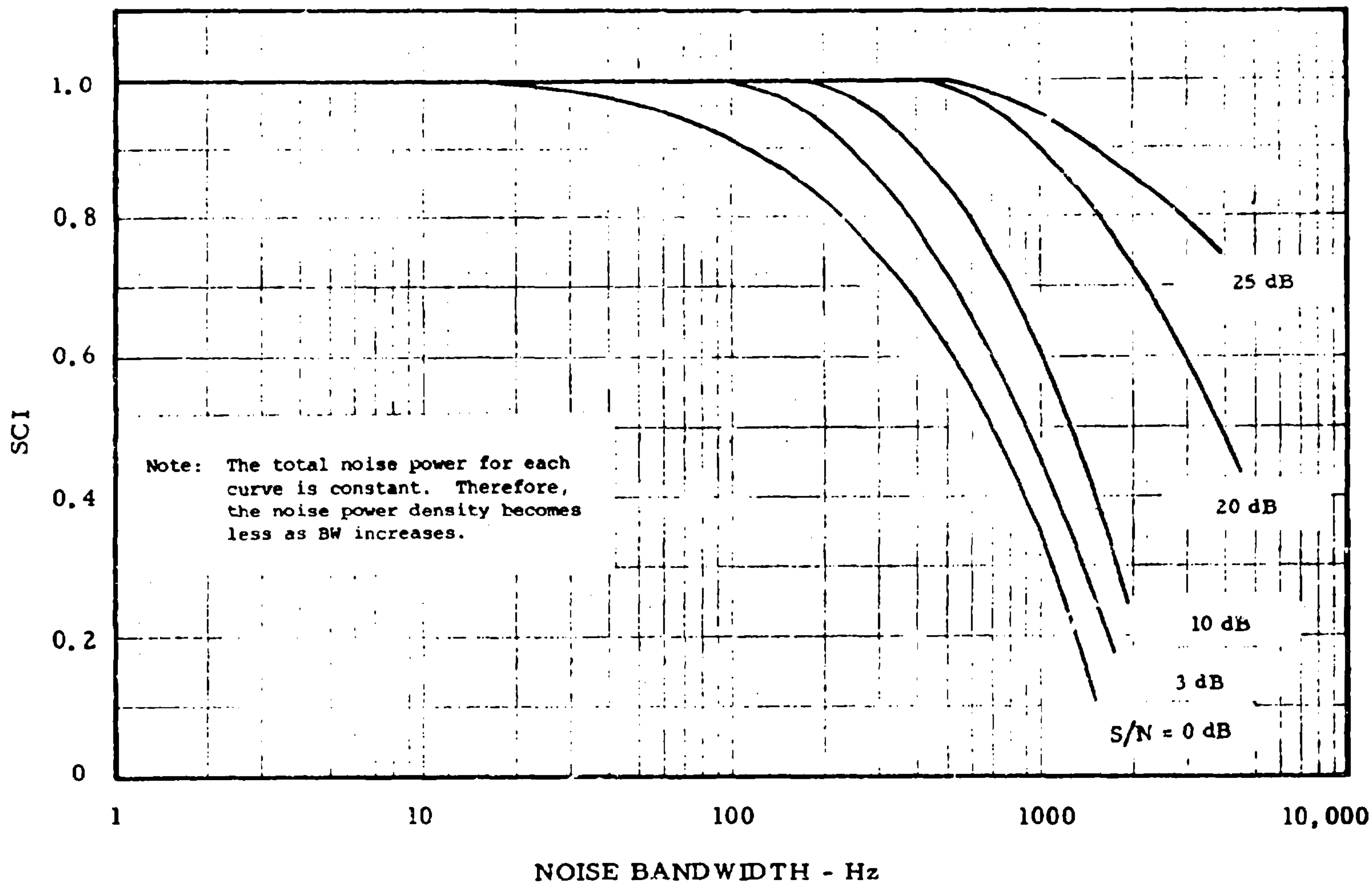


Figure 14. Effects of uniform density, Gaussian noise on SCIM measurements.

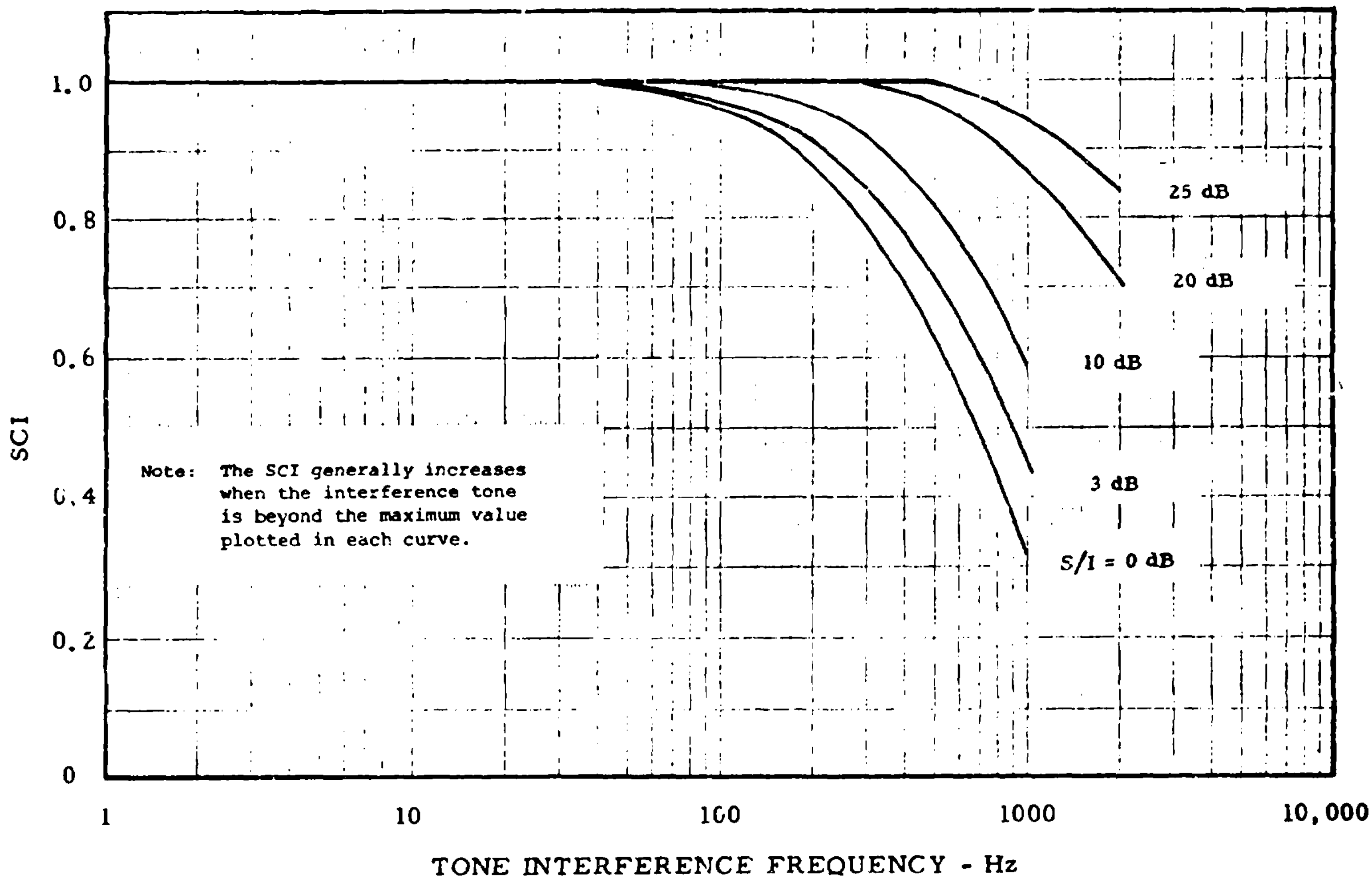


Figure 15. Effects of tone interference on SCIM measurements.

intelligibility. These data should be available in the near future. Figures 14 and 15 indicate the direct effects on the SCIM system, however.

4.2 Application to Digital Voice Systems

The laboratory evaluation of the SCIM technique for measuring AI was extended from the analog voice systems to digital voice processes. Two particular schemes known as pulse code modulation (PCM) and delta-modulation (DM) were of interest, since they are the most prevalently used for the transmission of voice in a digital format. Based on a priori knowledge of these digital processes and their inherent quantizing error (or noise), it was felt that a scoring technique such as SCIM would suffer inaccuracies when used in a digital voice channel. The initial tests performed were done with the use of simulation techniques in a digital computer. Programs were written for both a PCM and a DM process in a general purpose computer. The SCIM test signal was digitized at a high sample rate, and subjected to the simulated PCM and DM processes in a CDC-3800 computer system. The output data were then analyzed for the AI score using the 20-band computer method reported by Hubbard and Payne [1974]. This analysis was performed with no additive noise or distortion, so that the result would reflect only the error caused by quantizing noise. A

significant low score was obtained in each case, confirming the expected result. It was originally planned to continue these investigations over a range of sampling rates in the simulation process. However, the total process was found to be too expensive due to the quantity of data needed for the simulation and AI programs.

A variable-slope delta-modulation coder and decoder system* was procured in the interim by ITS for application in other programs. This system was placed on loan for use in the work reported here, and additional tests were performed in the laboratory in a back-to-back mode. The on-line SCIM system was used for measurement, and the results are presented in figure 16.

Very poor SCIM measurements were obtained at the lower DM sampling rates, and as figure 16 illustrates the resulting scores were quite erratic up to sampling rates as high as 60 kHz. The general performance curve was found to be poorer than expected based only on quantizing noise. Thus, spectral analysis was performed on both the "speech" and noise portions of the SCIM test signal during these measurements in order to find the cause. It was found that the "speech" signal did not suffer any significant change in

*KSMB Systems, Inc., Huntington, NY

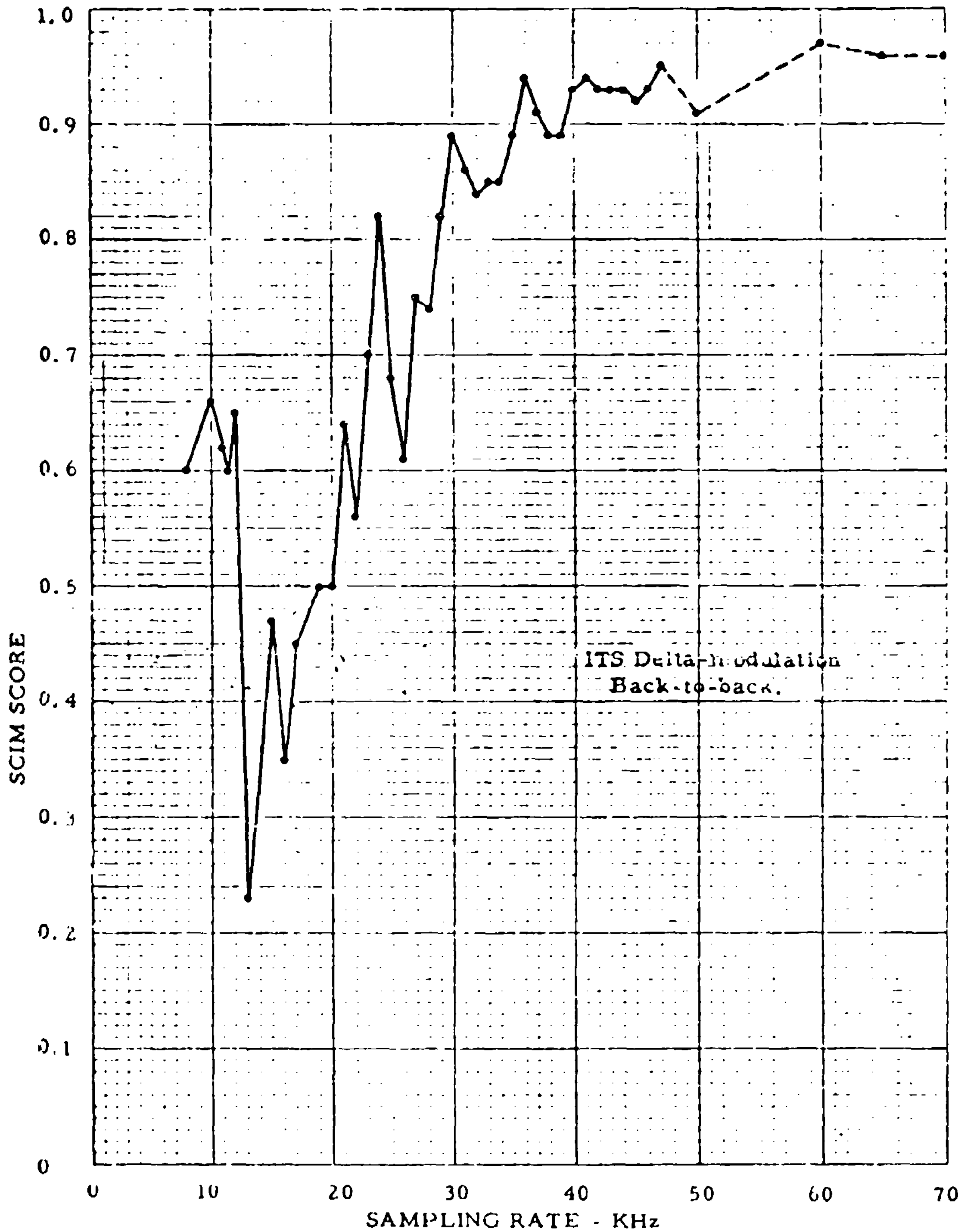


Figure 16. SCIM measurements made through the ITS delta-modulation system, back-to-back mode.

spectrum, but the noise spectrum changed radically with sampling rate. The spectra indicated that the problem is a combination of the quantizing noise and an interaction between the SCIM test tones and the sampling process during the noise portions of the signal. The noise density spectra at the lowest SCIM scores generally consisted of several sharp spikes somewhat uniformly spaced in frequency. At other data points, the spectra were more uniform in density and broad in frequency. As a part of Task B of the current project, we are experimenting with several methods to modify the test signal in order to alleviate the observed problem.

It was stated in the Introduction, that any objective technique should have a high degree of correlation with subjective performance measures to be most useful. Recent performance tests performed by NASA and objectively scored by ITS using the SCIM system have provided some comparative data for both analog and digital voice systems. The results of these tests were reported by Hubbard & Payne [1974], and are summarized in figure 17. NASA used the modified rhyme test (MRT) noted in figure 2, and performed both SCIM measurements and the MRT subjective evaluations for three voice systems. The standard curve for the MRT test is reproduced in figure 17 as the dashed-line curve. Note that the analog system, narrow-band frequency modulation (NBFM),

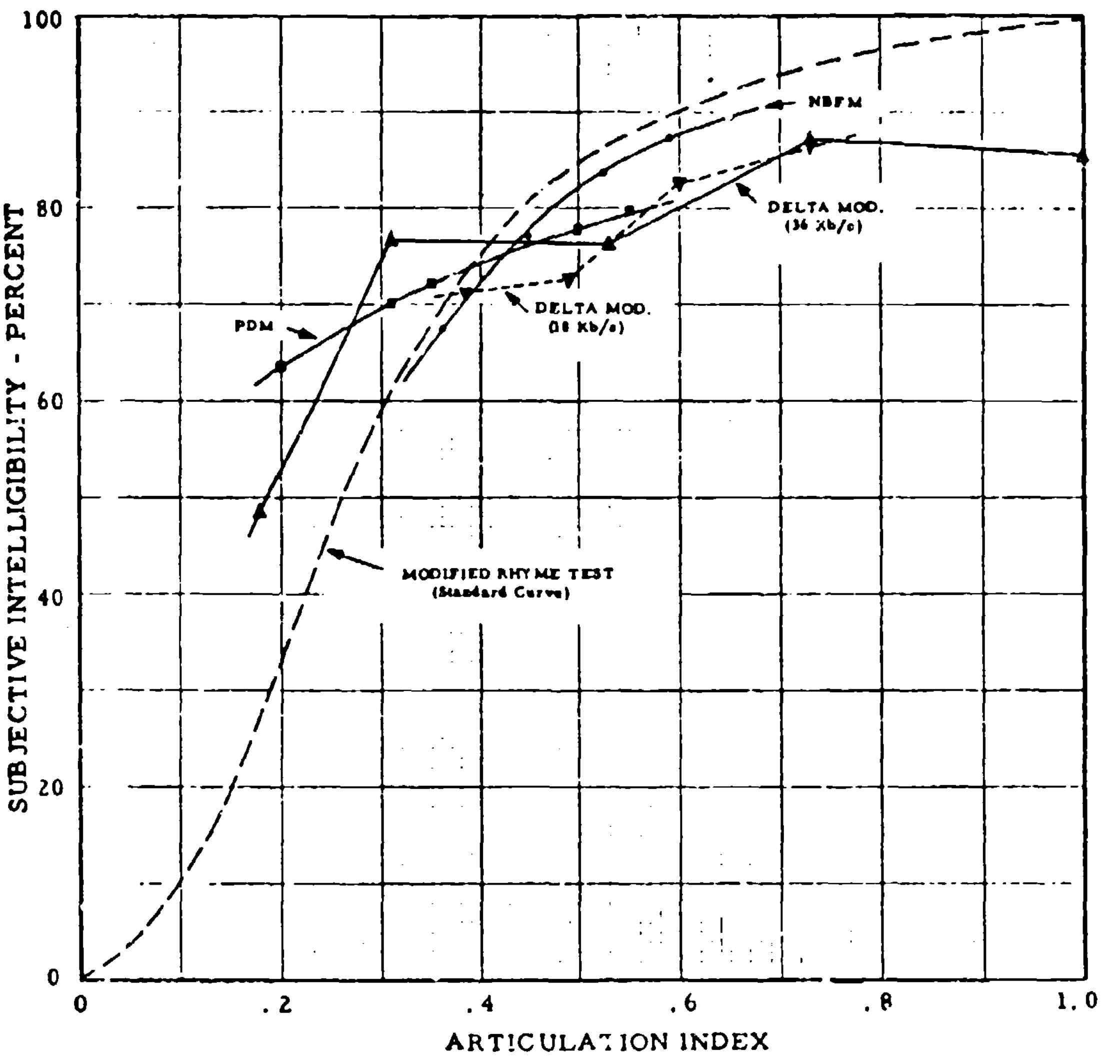


Figure 17. Comparison of articulation index measurements and subjective intelligibility scores for several voice communication systems.

gave measured results quite close to the standard curve and thus with the expected correlation. The digital system results, however, for pulse duration modulation (PDM) and for two sampling rates in a DM system, show poor correlation with the expected analog curve. These results coupled with that noted in figure 16 suggest that the objective measure implemented by SCIM is not an optimum method for digital channels. This particular problem is the subject for further study in the Task B effort of the project. In addition, it is felt that the work based on the LPC process mentioned in section 2 will be quite relevant to the digital voice problem.

5. SUMMARY

This report has presented the results of a study of objective techniques to measure the intelligibility performance of speech communication systems. The Articulation Index (AI) method is deemed to be the best available technique at this time, and the report discusses the instrumentation required for performing and analyzing system measurements.

Some of the limitations of the AI method were investigated in the study, and specific monitoring methods and correction data are presented to improve the accuracy of the measurement. Methods for measuring AI are based

exclusively on the automated method known as the Speech Communication Index Meter (SCIM). This system can be used to perform on-line measurements. The method is also outlined for use in an off-line mode where test data can be scored in general purpose computer routines.

Initial tests of the SCIM process in the digital modulated voice channel have pointed to some fundamental problems. These will be investigated further under the second phase of the continuing study.

6. ACKNOWLEDGEMENTS

The authors wish to thank Mr. Frank Kadi and Mr. Garth Kanen of the Systems Research and Development Service, Federal Aviation Administration for their continued interest and support in this study. We also particularly thank Mesrs. J.B. Wood and R.L. Ehret for their laboratory support, J.A. Payne and M.J. Miles for their computer programming assistance, and to Ms. Marilyn Longan for her preparation of the manuscript.

7. REFERENCES

ANSI (1969), Methods for the calculation of the articulation index, Standard No. S3.5-1969, Am. National Standards Inst., Inc., 1430 Broadway, NY, NY 10018.

Atal, B.S., and S.L. Hanauer (1971), Speech analysis and synthesis by linear prediction of the speech wave, J. of Acoustical Soc. of Am., 50, 637.

Busch, A.C. (1969), How do we get from theory to practice?, NAFEC Working Paper, presented at IEEE Intnl. Conf. Man/Machine Systems, St. John's College, Cambridge, England.

Dabbs, J.H. and O.L. Schmidt (1972), Appolo experience report-voice communications techniques and performance, NASA Technical Note TN D-6739.

Egan, J.P. (1948), Articulation testing methods, Laryngoscope, 58, 955-991.

French, N.R. and J.C. Steinberg (1947), Factors governing the intelligibility of speech sounds, J. Acoustical Soc. of Am., 19, 90-119.

Gierhart, G.D., R.W. Hubbard and D.V. Glen (1970), Electrospace planning and engineering for the air traffic environment, FAA Report No. FAA-RD-70-71, Systems Res. and Dev. Service, Washington, DC 20590.

Hubbard, R.W. and J.A. Payne (1974), Computer methods for the objective evaluation of speech communication systems, Office of Telecommunications, OT Report 74-29.

Kryter, K.D. (1962a), Methods for the calculation and use of the articulation index, J. Acoust. Soc. of Am., 34, 1689-1697.

Kryter, K.D. (1962b), Validation of the articulation index, J. Acoust. Soc. of Am., 34, 1698-1702.

Kryter, K.D. and J.H. Ball (1964), A meter for measuring the performance of speech communication systems, Tech. Doc. Rept. No. ESL-TDR-64-674, Electronic Systems

Division, Air Force Systems Command, L.G. Hanscom Field, Bedford, MA.

Licklider, J.C.R. (1946), The effects of amplitude distortion upon the intelligibility of speech, J. Acoust. Soc. of Am., 18, 429-434.

Lockheed Electronics Co. (1972), Briefing: The voice interference analysis sytem (VIAS), Lockheed Report No. A70-213, Tucson, Arizona.

Miller, G.A. and J.C.R. Licklider (1950), The intelligibility of interrupted speech, J. Acoust. Soc. of Am., 22, 167-173.

Stuckey, C.W. (1963), Investigation of the precision of an articulation-testing program, J. Acoustical Soc. of Am., 35, 1782-1787.

APPENDIX

COMPUTER METHODS OF AI SCORING

In order to provide other laboratories with a method for objectively scoring voice-system performance data, the ITS has developed computer programs for SCI and AI evaluations.

The first method developed is one in which the spectral analysis of the SCIM test signal is performed in a special purpose time-series analyzer, and the spectral data are then processed in a general purpose computer to obtain the SCI. The spectral analysis is performed in a Time/Data-100* system, operating sequentially on the "speech" and on noise portions of the signal over a spectral range of 5 kHz.

The Time/Data-100 accepts either analog or digital signals at its input, operates in a completely digital mode, and yields output data in either mode. For the SCI computations, the spectral information is recorded in digital format on magnetic tape, and these tapes are then scored in a general purpose computer. The computer program is written in Fortran IV, and is implemented

---

*Time/Data Corporation, 1050 E. Meadow Cr., Palo Alto, CA 94306.

in the ITS Laboratories on a CDC-3800 system. The flow diagram of the program is given in figure A1, and a complete listing is presented in table A1.

The SCIM test tapes are played in real-time into a Time/Data-100 time-series analyzer to obtain a spectral analysis. The TD-100 has an internal A/D converter set for a sampling period of 100 μsec (sampling rate of 10 kHz). Computation algorithms in this system are block oriented, with 1000 digital words per block. Thus, one block of input data represents a sample 0.1 sec in length. The spectral-average algorithm is selected, and a total of ten 0.1-sec blocks of data are processed during each 1 sec (nominal) period of the SCIM signal. These blocks are averaged, and the spectral density function is computed over a bandwidth of 5 kHz. The output data consist of 1001 12-bit data words with a resolution of 5 Hz/word. These data are recorded on digital magnetic tape for further processing. The general approach should be adaptable to any laboratory that has ready access to a spectral analysis system and a general purpose computer to format the SCI computation.

The techniques employed in this process were patterned after the SCIM system, in which 9 frequency bands are analysed and the spread-of-masking is handled in much the same manner as that used in the SCIM analyzer. For example,

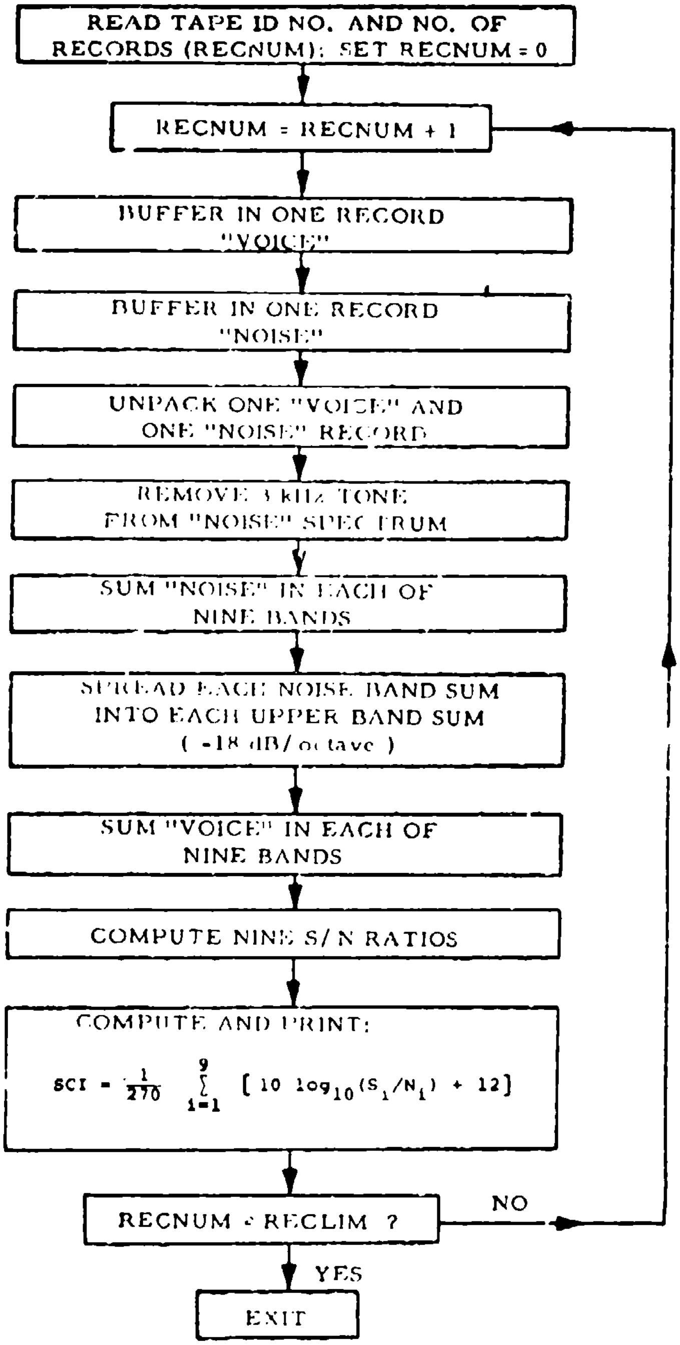


Figure A1. Flow diagram of SCIM program.

Table A1. Listing of 9-band SCIM computer program.

```
      PROGRAM SCIM3800
C-THIS PROGRAM COMPUTES SPEECH COMMUNICATION INDEX FROM MAG. TAPES OF
C-SIMULATED VOICE AND NOISE SPECTRA COMPUTED BY TD-100.
      DIMENSION SNOISE(1001),PACKNSU(251)
      DIMENSION TN(251),TV(251)
      DIMENSION ID(2), IDENT(2), PACKVC(251), PACKNS(251), DECBEL(9)
      DIMENSION VOICE(1004), NOISE(1004), UNPKVC(4)
      DIMENSION UNPKNS(4), FREQ(10), VCBAND(9), NSBAND(9), RATIO(9)
      DIMENSION SPRV(20),SPRN(20)
      INTEGER RECLIM, RECNUM, PACKVC, PACKNS, UNPKVC, UNPKNS,PACKNSU
      INTEGER FREQ, FUL, FLL,                  RECNM1
      TYPE REAL NOISE,NSBAND,MI
      DATA((FREQ(I),I=1,10)=250,500,750,1000,1250,1650,2050,2450,
     13200,4200)
      IDIF=0
      NEOF=0
      BW=0.20
      OKC=1000.*BW
      DO 772 I=1,10
      A=FREQ(I)*BW
  772 FREQ(I)=A+1
C  MULTPLYING BY BW CAUSES THE FREQ TO AGREE WITH THE ARRAY OF 1001 DENSITY
C  BW CNTD- VALUES OF SPECTRUM (IE DEPENDS ON THE SR)
      IF(FREQ(10).GT.1000)FREQ(10)=1000
  777 FORMAT(1H01018)
      PRINT 777,FREQ
      READ(60, 1) (IDENT(I), I = 1, 2), RECLIM,NFILSKP,FLAG,IDIF
C DATA CARD EXPLANATION
C  NFILSKP-NUMBER OF FILES TO SKIP ON LU 40
C  RECLIM STOPS THE PROGRAM AFTER RECLIM NUMBER OF SIGNALS HAVE PROCESSED
C IDIF - NUMBER OF SPECTRAL VALUES TO BE SHIFTED IN THE SPECTRUM.
C  IDIF CONTD- (IE TO SHIFT UP 100 HZ WHEN S R=10K, SET IDIF=-20)
C  SET FLAG=0, WHEN PROGRAM FINDS THE SPECTRUM SHIFT,1RST 3 RECS MUST BE
C  FLAG CONTD-1000 HZ TONE.
C SET FLAG =0 WHEN LU 40 BEGINS WITH 3 CONSECUTIVE I KC TONE RECS
C  SET FLAG=1  WHEN  NO TEST FOR FREQ SHIFT IS USED.
    1 FORMAT (2A8, 118,16,F5.0,15)
      DO 38 I=1,NFILSKP
38    CALL SKIPFILE (40)
      PRINT 37,NFILSKP
 37   FORMAT( *    SKIPPED *14,*  FILES ON 40*)
      RECNUM = 0
C     GO TO 100
   40 RECNUM = RECNUM + 1
   41 BUFFER IN (40, 1) (PACKVC(1), PACKVC(251))
   42 IF(UNIT, 40) 42, 70, 50, 60
   50 PRINT 51, RECNUM
   51 FORMAT (10X, *EOF OCCURRED WHILE READING RECORD NO. *, I6)
      NEOF=NEOF+1
      PRINT 52,NEOF
52     FORMAT(*1   END OF RUN*I5)
      IF(NEOF.GE.2)CALL EXIT
C     IF(NEOF.GE.5)CALL EXIT
      RECNUM=0
```

Table A1. (Continued)

```
      GO TO 40
   60 PRINT 61, RECNUM
   61 FORMAT (10X, *PARITY ERROR OCCURRED WHILE READING RECORD NO. *,I6)
70    CONTINUE
   71 BUFFER IN (40, 1) (PACKNS(1), PACKNS(251))
   72 IF(UNIT, 40) 72, 92,  80, 90
   80 PRINT 81, RECNUM
   81 FORMAT (10X, *EOF OCCURRED WHILE READING RECORD NO. *, I6)
      GO TO 200
   90 PRINT 91, RECNUM
   91 FORMAT (10X, *PARITY ERROR OCCURRED WHILE READING RECORD NO. *,I6)
C-UNPACKING 1 VOICE RECORD AND 1 NOISE RECORD INTO TWO 1001 ELEMENT ARRAYS.
   92 CONTINUE
   93 BUFFER IN(40,1)(PACKNSU(1),PACKNSU(251))
   94 IF(UNIT,40)94,101,95,96
   95 PRINT 81,RECNUM
      GO TO 200
   96 PRINT 91,RECNUM
  101 CONTINUE
      J=FREQ(6)/4
      J=9
      DO 97 I=J  ,251
   97 PACKNS(I)=PACKNSU(I)
  100 DO 106 J = 1, 251
      DECODE(8, 105, PACKVC(J)) UNPKVC
      DECODE(8, 105, PACKNS(J)) UNPKNS
  105 FORMAT(4R2)
      JM1X4 = (J - 1) * 4
      DO 106 I = 1, 4
      JM1X4I = JM1X4 + I
      NOISE(JM1X4I) =      UNPKNS(I)
      VOICE(JM1X4I) =      UNPKVC(I)
       I1=JM1X4I
  106 CONTINUE
      DO 309 I=1,101
      TV(I)=VOICE(I)
  309 TN(I)=NOISE(I)
      IC=1
      DO 310 I=2,1001
      JV=(I-1)/10
      JV=JV* 10
      VOICE(I)=0
      NOISE(I)=0
      IF(JV.EQ.I-1)IC=IC+1
      IF(JV.EQ.I-1)VOICE(I)=TV(IC)
      IF(JV.EQ.I-1)NOISE(I)=TN(IC)
  310 CONTINUE
C     EXAMINE TONE FOR FREQUENCY SHIFT
C SET FLAG TO ZERO WHEN INPUT TAPE BEGINS WITH THREE CONSEC RECORDS OF 1 KC
      IF(FLAG.EQ.1.)GO TO 404
      FLAG=1.
      AMAG=500.
      ILW=1
      DO 401 I=100,500
      IF(VOICE(I).LE.AMAG) GO TO 401
      ILAST=I
      IF(VOICE(ILAST).GT.VOICE(ILW))ILW=ILAST
401   CONTINUE
      IDIF=201-ILW
```

Table Al. (Continued)

```
      PRINT 201,'VOICE
 404   CONTINUE
      PRINT 207,IDIF
207    FORMAT(*  DATA WAS SHIFTED *I5,*  PLACES*)
      ILAS=1001+IDIF
       IF(ILAS.GT.1001)ILAS=1001
      DO 402 I=1,ILAS
      IV=I-IDIF
      IF(IV.LE.0)GO TO 402
      IF(IV.LT.I)GO TO 403
      VOICE(I )=VOICE(IV)
      NOISE(I )=NOISE(IV)
      GO TO 402
C  SHIFT TO RIGHT
403   IL=1002-IV
      IB=1002-I
      VOICE(IL)=VOICE(IB)
      NOISE(IL)=NOISE(IB)
402   CONTINUE
C     PRINT 201, VOICE
C     PRINT 201, NOISE
      IF(RECNUM.EQ.0) GO TO 40
400   CONTINUE
  201 FORMAT (1X, 20F6.1)
      PRINT 202
  202 FORMAT (* *)
  204 FORMAT(10F10.0)
      PRINT 202
   26 CONTINUE
      DO 121 I=1,1001
      VOICE(I)=VOICE(I)-NOISE(I)
      IF(VOICE(I).LT.0.)VOICE(I)=0.
121   CONTINUE
C-SUMMING 9 NOISE BANDS.
      DO 126 I = 1, 9
      NSBAND(I) = 0
      VCBAND(I)=0.
      FLL = FREQ(I)
      FUL = FREQ(I + 1)
      LN=0
      SUMP=0
      DO 124 J = FLL, FUL
      IF((NOISE(J).LE.0.).AND.(VOICE(J).LE.0.))GO TO 124
      IF(NOISE(J).LE.0.)GO TO 123
      RTO=VOICE(J)/NOISE(J)
       GO TO 122
C  AT STATEMENT 123 ONE HAS A VALUE FOR VOICE AND A ZERO FOR NOISE
123   RTO=VOICE(J)
122    CONTINUE
      LN=LN+1
      SUMP=SUMP+RTO
      NSBAND(I) = NSBAND(I) + NOISE(J)
124    CONTINUE
      IF(LN.EQ.0)GO TO 126
      SUMP=SUMP/LN
      VCBAND(I)=NSBAND(I)*SUMP
      IF(NSBAND(I).EQ.0.)VCBAND(I)=SUMP*LN
126   CONTINUE
      PRINT 204, NSBAND
      PRINT 204, VCBAND
      PRINT 202
```

Table A1. (Continued)

```
C-SPREADING NOISE UPWARD AT -18DB/OCT. INTO 9 NOISE BANDS.
      DO 24 I=1,9
      DO 12 J=1,9
      IF(FREQ(I).LE.FREQ(J+1))GO TO 12
C     AL=((FREQ(I)-FREQ(J+1))/(0.5*FREQ(J+1)))*1.8
      AL=((FREQ(I)-FREQ(J+1))/(    FREQ(J+1)))*1.8
      PER=10**AL
      YI=NSBAND(J)/PER
      IF(YI.GE.NSBAND(I))NSBAND(I)=YI
      SPRN(J)=YI
      YI=VCBAND(J)/PER
      SPRV(J)=YI
   12 CONTINUE
C     DO 14 J=1,9
C     IF(NSBAND(J).LT.SPRN(J))NSBAND(J)=SPRN(J)
C     IF(VCBAND(J).LT.SPRV(J))VCBAND(J)=SPRV(J)
C14    CONTINUE
   24 CONTINUE
      PRINT 204, NSBAND
      PRINT 202
C-SUMMING 9 VOICE BANDS.
      PRINT 204, VCBAND
      PRINT 202
C-CALCULATING 9 S/N RATIOS.
      DO 150 I = 1, 9
      IF(NSBAND(I) .LE. 0 .AND. VCBAND(I) .GT. 0) GO TO 142
      IF(NSBAND(I) .LE. 0 .AND. VCBAND(I) .LE. 0) GO TO 144
      GO TO 146
  142 RATIO(I) = 63.1**2
C 142 RATIO(I) = 63.1
      GO TO 150
  144 RATIO(I) = 1./4.
      GO TO 150
  146 FVCBND = VCBAND(I)
      FNSBND = NSBAND(I)
      RATIO(I) = FVCBND / FNSBND
C     RATIO(I)=SQRTF(RATIO(I))
  150 CONTINUE
      PRINT 203, RATIO
  203 FORMAT (1X, 10E13.6)
      PRINT 202
C-CONVERTING 9 S/N RATIOS TO DECIBELS.
      DO 160 I = 1, 9
C-DIFFERENCE BETWEEN LONG TERM RMS AND SPEECH PEAKS =  6DB = RATIO OF 4.
C     IF(RATIO(I) .LE.1.) RATIO(I) = 1./50.+1.
      IF(RATIO(I).LE. 0.) RATIO(I) = 1./50.
C     DECBEL(I) =10.*ALOG10( 2*(RATIO(I)-1))+12.
      DECBEL(I) =10.*ALOG10(    RATIO(I)    )+12.
      IF(DECBEL(I).LE.0. )DECBEL(I)=0.
      IF(DECBEL(I).GT.30.)DECBEL(I)=30.
  160 CONTINUE
      PRINT 203, DECBEL
      PRINT 202
C-CALCULATING SPEECH COMMUNICATION INDEX.
```

Table Ai. (Continued)

```
      SUM = 0
      DO 170 I = 1, 9
  170 SUM = SUM + DECBEL(I)
      SCI=SUM/270.
      RECNM1 = RECNUM - 2
      PRINT 180,         RECNUM, SCI
  180 FORMAT (5X, *SPEECH COMMUNICATION INDEX FOR TRIAL   NUMBER *,I5,
     A  7X,F6.3)
      IF(RECNUM .GE. RECLIM) GO TO 200
      GO TO 40
  200 CALL EXIT
      END
```

figure A2 shows the spectral division of the noise present in SCIM analysis, in which the noise power is spread upwards in the spectrum at a slope of -18 dB/octave. Downward spread is negligible in the SCIM computation and the 60 dB/octave skirt shown for the low side of the curves in figure A2 represents the best achievable for the analog filter design. These same characteristics are produced in the computer program by appropriate weighting of the input spectral data.

This scoring method was originally developed for use in a program sponsored by the FAA [Hubbard, R.W., D.V. Glen, and W.J. Hartman (1970), Modulation characteristics critical to frequency planning for the aeronautical services, unpublished ESSA Tech Memo ERLTM-ITS232, (April)] and applied to the analysis of Air Traffic Control (ATC) system measurements [Gierhart et al., 1970]. The method was compared directly with a calibrated set of SCIM measurements for verification. The results are shown in table A2.

The TD-100/CDC-3800 method was also applied to the analysis of system performance measurements made by NASA/GSFC. Results of these measurements are presented in a recent report by Hubbard and Payne [1974].

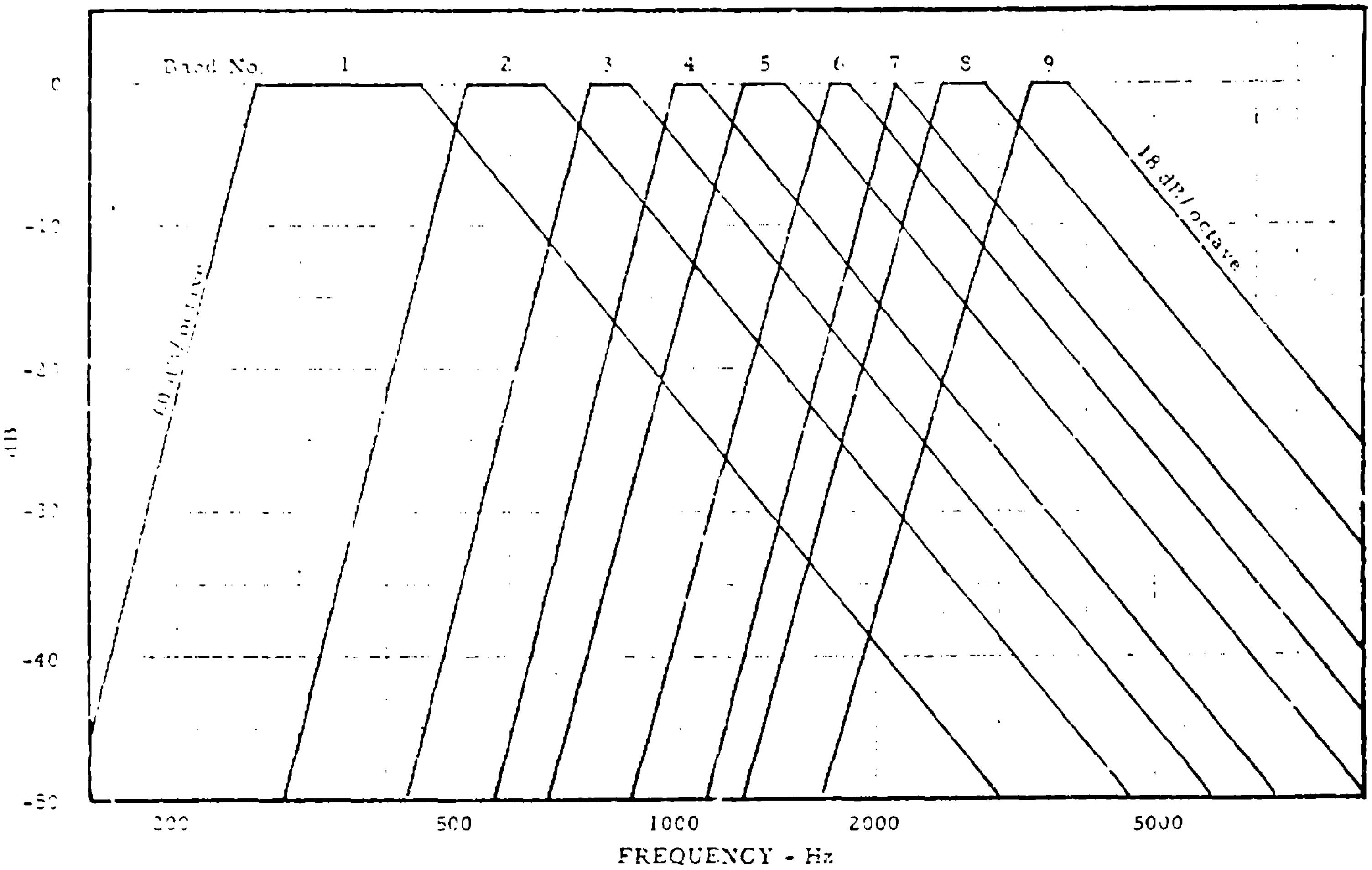


Figure A2. Spectral division of the signal and noise in SCIM analysis.

Table A2. Calibration of TD-100/CDC-3800 program

SCIM Reading (10 samples)	Average Program (10 samples)	Standard Deviation
0.36	0.32	.02
0.52	0.51	.04
0.75	0.74	.03
0.99*	1.00	.00

*The SCIM analyzer presents a maximum reading of 0.99 for a perfect score.

The second computer method was developed under sponsorship of NASA/GSFC. The objective of the project was to develop a complete program for the computation of AI based on the full 20-band method specified by the ANSI [1969]. This objective was accomplished in the ITS Laboratories, implementing the program on a CDC-3800 system and verifying the results with the same calibration values listed in table A2. The results were essentially the same as those found for the TD-100/CDC-3800 method. The program was furnished to NASA, and it has been successfully

implemented on an IBM-360/95 system for application to test programs conducted by the Applications Experiments Branch.

The AI computer program includes the spectral analysis of the test signal, using Fortran subroutines that include a fast-Fourier transform (FFT) algorithm. The AI scoring program is very similar to that described above for the SCI, with the exception that the full 20-band analysis method is used. A detailed description of the program logic and the CDC-3800 listing is presented in Hubbard & Payne [1974]. Application of this program to test measurements requires analog-to-digital (A/D) conversion of the measured data. Details of the A/D process are included in the above reference.

REFERENCES

ANSI (1969), Methods for the calculation of the articulation index, Standard No. S3.5-1969, Am. National Standards Inst., Inc., 1430 Broadway, NY, NY 10018.

Gierhart, G.D, R.W. Hubbard and D.V. Glen (1970), Electrospace planning and engineering for the air traffic environment, FAA Report No. FAA-RD-70-71, Systems Res. and Dev. Service, Washington, DC 20590.

Hubbard, R.W. and J.A. Payne (1974), Computer methods for the objective evaluation of speech communication systems, Office of Telecommunications, OT Report 74-29.